

A FACTORIAL STUDY OF INTROVERSION-EXTRAVERSION

By H. P. HILDEBRAND

Tavistock Clinic

ABSTRACT

Previous research suggested that Jung's Theory of Psychological Types could best be examined by factorial methods. Two orthogonal factors of neuroticism and introversion-extraversion were hypothesized in order to explain Jung's theory. A varied neurotic population was examined on a large battery of tests, and factors identified as 'neuroticism', 'g' and 'introversion-extraversion' extracted. Further evidence confirming Jung's theories was given by using regression equations and factor scores which enabled criterion groups not included in the original matrix to be correctly identified.

I. INTRODUCTION: DEFINITION

Jung (1913) is commonly acknowledged to have coined the terms 'Introversion' and 'Extraversion' to describe a typology which he traces back to its earliest beginnings—the well-known division into 'habitus pythisicus' and 'habitus apoplecticus' elaborated by Galen. Other modern authors before Jung, notably Jordan (1890) and Gross (1902), had put forward psychological theories based on a type division, but Jung was the first to gain widespread acceptance for the type theory. Since his descriptions of the introvert and the extravert have been the basis for much of the subsequent work in this field, it would seem advisable to quote them in some detail.

Introversion: means a turning inward of the libido whereby a negative relation of subject to object is expressed. Interest does not move towards the object but recedes toward the subject. Everyone whose attitude is introverted thinks, feels and acts in a way that clearly demonstrates that the subject is the chief factor of motivation while the object at most receives only a secondary value. (Introversion may possess either a more intellectual or a more emotional character) just as it can be characterized by either intuition or sensation. Introversion is active when the subject *wills* a certain seclusion in the face of the object: it is passive when the subject is unable to restore again to the object the libido which is streaming back from it. When Introversion is habitual, one speaks of an *introverted type* (1924, p. 567).

Extraversion: means an outward turning of the libido. With this concept I denote a manifest relatedness of subject to object in the sense of a positive movement of subjective interest towards the object. Everyone in a state of extraversion thinks, feels and acts in relation to the object, and moreover in a direct and clearly observable fashion, so that no doubt can exist about his positive dependence upon the object. In a sense, therefore, extraversion is an outgoing transference of interest from the subject to the object. If it is an intellectual extraversion, the subject thinks himself into the object; if a feeling extraversion then the subject feels himself into the object. The state of extraversion means a strong, if not exclusive, determination by the object. One should speak of an *active* extraversion when deliberately willed, and of a *passive* extraversion when the object compels it, i.e. attracts the interest of the subject of its own accord, even against the latter's intention. Should the state of extraversion become habitual the *extraverted type* appears (1924, p. 543).

Three further points are of importance. First, Jung states that

although there are individuals in whom one can recognize the type at first glance, this is by no means always the case. As a rule only careful observation and weighing of the evidence permits a sure classification. However clear and simple the fundamental principle of the opposing attitudes may be, this concrete reality is none the less complicated and obscure, for every individual is an exception to the rule. Therefore, one can never give a description of a type, no matter how complete, which absolutely applies to one individual, despite the fact that thousands might, in a sense, be characterised by it.

Secondly, following the implications that I-E is a continuum, Jung identifies the poles of the distribution of I-E with two of the main groups of neurotic disorders—the extravert with the conversion hysterics, and the introvert with Janet's psychasthenias. He says (1924, p. 421); 'Hysteria is in my view, by far the most frequent neurosis with the extraverted type', while psychasthenia is identified as the typical neurotic disorder of the introvert. This neurosis is characterized on the one hand by marked sensitivity, and on the other by great exhaustion and constant tiredness. Eysenck has suggested the term 'dysthymic' to cover this syndrome of affective disorders. In modern psychiatric terminology dysthymic would cover such syndromes as anxiety states and reactive depressions.

Thirdly, Jung stresses the independence of introversion and neurosis: he says—'It is a mistake to believe that introversion is more or less the same as neurosis. As concepts the two have not the slightest connexion with one another.'

Jung's formulation was widely accepted and a great number of studies carried out in an attempt to provide an experimental basis for the theory. But as Collier & Emch (1938) point out, there was considerable confusion as to the definition of the traits it was hoped to define. The main lines of Jung's arguments were ignored, while the terms introversion and extraversion were retained and redefined according to the authors' theoretical preferences. Collier and Emch remark that there are three main varieties of redefinitions which still claim to have some connexion with Jung's original dichotomy. First, there is the attempt at simplification; secondly, the identification of introversion with neurotic tendency—extraversion being given 'per contra' a connotation of normality; and thirdly, the relating of introversion to the schizophrenic and extraversion to the manic depressive psychoses.

Not surprisingly the experimental evidence is very inconclusive. Experiments based on the first variety of redefinitions are quite inadequate to explain the wide variety of phenomena cited by Jung. The second type of definition is based on Freud's adaptation of the term introversion. For Freud, introversion is 'one of the invariable and indispensable considerations in every case of psycho-neurosis—a substitution for actual objects of phantasies of these objects' (Freud, 1920). It is Freud's belief that Jung's use of the term is 'indiscriminate'. As Collier and Emch point out 'Freud's definition of the term represents a rigid contraction both in meaning and in use'. They maintain that Freud's definition of introversion has been particularly pervasive in the work of writers interested in the subject, but paradoxically the writers concerned seem to regard it as a natural extension of Jung's concept rather than a contraction. It is beyond the scope of this paper to examine this aspect of the controversy, but it is apparent that any work based on Freud's use of the term introversion can have little relation to the original formulation. Similarly, of course, any attempt to identify introversion with psychotic manifestations goes far beyond Jung's formulations, and can be disregarded in terms of validation of the hypothesis.

There seems no point in carrying the argument further, particularly as all early experimental work is summed in three excellent reviews of work carried out to establish the empirical basis of I-E. They are those of Guilford (1934), Collier & Emch (1938), and Eysenck (1947), and are devoted in the main to small-scale experiments. Two further causes of confusion presented by these authors are, first, the clinical identification of introversion with neurosis referred to above, and secondly, the identification of extra-

version with sociability first put forward by Freyd (1924). This identification was disproved when put to an empirical test by Russell Frazer (1947). Lack of sociability must be regarded as an index of neuroticism, not as a sign of introversion (Eysenck, 1947).

Eysenck (1953*a*) has also presented a review of factorial studies with some bearing on Jung's hypothesis. In the main these studies agreed in finding 'a bi-polar factor variously called "surgency-desurgency" (Cattell, 1946), "c" (Garrett, 1921), "aggressive-inhibitive" (Line, 1935) or whatever terms seem appropriate... these factors closely agree with one another as well as our own hysteria-dysthymia factor'. All the studies named above are based upon rating scales and have their origin in the pioneer work of Heymans & Wiersma (1906) and of Webb (1915). The name given to a factor arising from the inter-correlation of a number of traits seemed mainly to depend on the theoretical bias of the author. The low reliability and validity of such studies make caution necessary when identifying such factors with Jung's Introversion-Extraversion hypothesis, but the similarity of the results is striking.

Further evidence is available from questionnaire studies. Of these the questionnaire constructed by Guilford is outstanding in terms of reliability and soundness of statistical construction. Guilford himself would explain I-E in terms of a number of oblique factors, but Lovell (1945) and North (1948) have shown that the questionnaire in fact gives rise to two second-order factors, the second of which shows a distinct resemblance to the factors found by the rating method (Hildebrand, 1956).

(i) *Rationale*

II. THE PRESENT STUDY

In the brief summary of previous work reported above, it has been shown that while small-scale testing has generally been ineffective to test Jung's hypotheses, various factorial studies have reported factors which seemed to resemble I-E. The evidence in favour of such a factor is suggestive but not convincing. In the present study, however, it was decided to test Jung's hypothesis directly, using factorial methods and a battery of tests designed to describe introversion-extraversion as completely as possible. The following hypotheses were derived from Jung's original statement by the writer. *Hypothesis 1*: That I-E (Introversion-Extraversion) is an underlying dimension of personality, common in some degree to all men. *Hypothesis 2*: That I-E is distinct from a second underlying dimension, namely, Neuroticism. *Deduction 2. 1*: That if a factor analysis is carried out on a group of tests purporting to measure I-E and Neuroticism two general factors should emerge, which can most meaningfully be explained as I-E and Neuroticism. *Deduction 2. 2*: That these two factors would be significant and orthogonal. *Hypothesis 3*: That extraverts would tend as a group to suffer from a hysterical type of neurosis; that introverts as a group would tend to suffer from an anxiety neurosis. *Deduction 3. 1*: That hysterics as a group would tend to score high on the Neuroticism factor and also tend to cluster at the E pole of the I-E factor; conversely although anxiety states as a group would also tend to score high on the Neuroticism factor they would tend to cluster at the I pole of the I-E factor. *Deduction 3. 2*: That mixed neurotic groups would score high on the Neuroticism factor, but would tend towards a central distribution on the I-E factor. *Deduction 3. 3*: That normal subjects would have a central tendency on the I-E factor and low scores on the Neuroticism factor. The deductions made from these hypotheses are available to proof or disproof by experimental investigation.

In view of the nature of the investigation it was decided to use so-called objective behaviour tests. ('Objective behaviour tests', as here used, implies only that the tests are scored objectively and not that the tests themselves are objective. Such tests are scored by a standard set of instructions which do not require subjective interpretation by the interpreter.) The reasons for this decision were that the use of such tests would enable the experiment to be repeated independently if required, that the reliability and validity of the tests were usually known and that no evidence was at that time available about the interaction of these tests with regard to the factors under investigation.

According to the criteria laid down above, two groups of tests were chosen, one of which had been shown by previous experiments to differentiate between groups of normals and neurotics, and a second which had been shown to differentiate significantly between dysthymics and hysterics in previous investigations. It was, of course, necessary that such differentiation did not depend upon subjective judgments but was a function of the test scores themselves. A group of tests with high loadings on the third factor 'g' was also included as previous work had suggested that there was some relationship between I-E and intelligence (Eysenck, 1947).

It was proposed to resolve the crucial point of the generality of the I-E factor in the following manner. If the writer's interpretation of Jung's hypothesis was correct, then the exclusion of the two groups, which he describes as being the most typical of the poles of the dimension, from a factorial experiment should not affect the extraction of a factor which would be most meaningful psychologically when described as I-E. The hysteric and anxiety groups would be treated as criterion groups and not included in the factor analysis. Similarly, the normal control group was regarded as a criterion group for the neuroticism factor.

After the analysis had been carried out it was further proposed to make estimations of the regression coefficients of each variable, and to plot the groups in a two-factor space with regard to the two factors of Neuroticism and I-E. If the anxiety states and hysterics were found to be significantly different on the I-E factor after their score had been estimated and not significantly different on the Neuroticism factor then deduction 3. 1 would be proved. A plot of the mixed neurotic groups would afford proof or disproof of Deduction 3. 2 and a plot of the normals would do the same for Deduction 3. 3.

An additional strength of this method is that it would confirm or disprove differentiation between criterion groups by means of analysis of variance techniques, and indicate where such differentiation was due to specific causes or the action of the I-E factor. It was also decided to include subjects suffering from all types of neurotic disorder, in order that the different syndromes could be related to the relevant personality dimensions of Neuroticism and I-E. A normal control group was also tested. It should be noted that the results found are those from a neurotic population and cannot immediately be generalized to cover the behaviour of a normal population.

(ii) *Subjects*

All neurotic subjects were drawn from in-patients at Belmont Hospital, Sutton, Surrey. All subjects were male. Minski gives the following details (1950): 'Approximately 1500 to 2000 patients are admitted to Belmont annually for an average length of stay of 2 to 3 months. By far the majority of admissions are neurotics and come from all over the country. All patients are voluntary.'

All admissions over an 8-month period were screened. From these 145 patients were selected as falling within the following categories. (1) Over 18 and under 50 years of age. (2) Considered by psychiatrist

to show no sign of organic lesion. (3) Considered by psychiatrist to show no sign of psychotic illness. (4) I.Q. above 90. (5) Not undergoing physical treatment (E.C.T. or Insulin Therapy). All patients were severe neurotics, having been given up by both G.P. and out-patient clinics, and the amount of mixed symptomatology was great. Cases where the psychiatrist felt justified in making a clear-cut diagnosis were few, and the groups into which the patients were sorted should be regarded as classifications in terms of outstanding syndromes.

Details of the breakdown of subjects into groups are given in Table 1. The total number of patients tested was 145. Overall analyses of variance between the groups showed no significant difference in age. As a group, however, the Obsessionals and Depressives tended to be somewhat older. The mean age was 33.9 years, with a standard deviation of ± 8.1 .

Table 1

Group	No.	Abbreviation
Conversion hysterics	25	Dys.
Psychopaths	20	P.P.
Reactive depression	10	Dep.
Obsessional neurosis	10	Obs.
Anxiety states	25	Anx.
Mixed symptomatology	55	Mx.

The normal control group consisted of 25 male soldiers taken at random from a reallocation centre in the neighbourhood. Eysenck states of a preceding and similar group to that tested by the writer: 'It is known that soldiers going through such a centre tend to have a higher degree of neurotic and psychotic abnormality than do soldiers not reallocated. In addition these soldiers tend to be slightly below average in intelligence, and therefore constitute a distinctly poor sample of the normal population.' That it is possible, in spite of this difficulty, to obtain highly significant differentiation between such a group and abnormal groups has been shown already (Eysenck, 1953*b*). The mean age of the Normal group equals 24.5 ± 6.6 years. The c.r. between the mean age of this group and the mean age of the neurotic group is significant (c.r. = 5.53). However, the difficulty of obtaining a homogeneous normal sample imposed the use of this group. The testing of the control group was carried out at the Maudsley Hospital, and the same procedure was used as with the neurotic groups at Belmont.

(iii) *Experimental variables*

In order to test the hypotheses propounded above a review was made of the literature and three groups of tests selected. The first of these was a group of three intelligence tests. Eysenck had suggested that intelligence might have some bearing on I-E and this group of tests was selected as covering the cognitive field. The tests were the Progressive Matrices, the Mill Hill Vocabulary and the Nufferno Power Test. Detailed descriptions of all tests and the rationale for use may be found in the writer's Ph.D. thesis (Hildebrand, 1953).

The second group of tests consisted of tests which it was hoped would define the Neuroticism factor. They were as follows: (4) The Maudsley Medical Questionnaire. (5) The Maudsley Medical Questionnaire—lie scale. (6) An annoyances Questionnaire—a list of thirty possible annoyances adapted from Pressey's X-0 test. Previous research suggested that neurotics would be more irritable than normal groups.

The third group of tests included those which previous workers had suggested might have some bearing on the I-E dimension. They were (7) A Humour Scale—subjects were presented with thirty cartoons and asked to classify them according to preference. There were ten cartoons each for Sex humour, Incongruous humour, and Extra-punitive humour. (8) The Guilford Martin Inventory of Factors S T D C R. Factors S and R had been found by previous workers to have some bearing on I-E, and factors T D and C on neuroticism (Hildebrand, 1956). (9) Static ataxia and (10) Suggestibility: four scores each were taken for static ataxia and for suggestibility to autosuggestion—maximum forward sway, maximum backward sway; total sway in half-inch units; and number of reversals (oscillations) in quarter-inch units. (11) The Porteus maze Test—adapted Q score. Scoring categories were Starting time, Tracing time, Weighted wrong directions, Lifted pencils, Crossed lines. (12) Ratio stature/Transverse chest diameter in inches. (13) Leg persistence—the time in which the subject holds his leg above a chair. Analyses of variance were calculated for all variables, in order to test hypotheses relating to the ability of the tests to

differentiate between neurotic groups. Critical ratios between means of the neurotic groups and that of the normal control group were also calculated where this seemed advisable. Full details will be found in Hildebrand (1953).

II. RESULTS

(i) *Statistical analysis.* Selection of variables for the final factor analysis was based on the following criteria. (1) That all subjects should have completed the test; (2) That the score of the total population on the test should have a normal distribution or a close approximation to it. It will be recalled that it was proposed to exclude the criterion groups from the matrix in order to provide external criteria of I-E and Neuroticism, and the population included in the final matrix was thereby reduced to 95 subjects and consisted of 55 mixed neurotics, 10 obsessionals, 10 reactive depressives and 20 psychopaths. There were in all 22 variables that satisfied the criteria. All scores of the population on the 22 selected variables were transformed into standard scores on a nine point scale. Product moment correlations were then calculated for all variables.

In view of the requirements of the hypotheses and the lack of satisfactory tests of significance which might be applied to the ordinary centroid method of factor analysis (Burt, 1952), it was decided to use Lawley's Maximum Likelihood Method (Thomson, 1950). This technique has the great advantage of applying what Burt calls the best test of significance of extracted factors, namely, a derivation of Bartlett's (1950) method. Lawley's method assumes specific factors; it aims at minimizing the residuals and so accounting for the intercorrelations to the greatest possible extent. 'The method has the property of using the largest amount of information available in the data and gives "efficient" estimates which roughly speaking, are on the average nearer the true values than those obtained by other "inefficient" methods of estimation' (Lawley, 1950). Lawley makes two initial assumptions before using this method. These are, first, that the test scores and the factors of which they are linear functions are approximately normally distributed; secondly, that a hypothesis must be set up as to the number of factors present. Both these assumptions have been met in the present study.

A full centroid analysis was performed on the data to give first estimates of factor loadings and three factors were extracted. The hypothesis was set up that these three factors, which were tentatively identified as 'Neuroticism', 'g' and 'I-E' were sufficient to account for the observed correlations. Two iterations by Lawley's method were carried out as recommended by Lawley and the adjusted factor loadings subtracted from the original matrix.

The resulting residual correlations were tested for significance by Bartlett's (1950) correction of Lawley's test of significance and it was found that there were still significant factors remaining in the matrix. A fourth factor was accordingly extracted. On the basis of this evidence it is suggested that all four extracted factors are significant. Two rotations were sufficient to adjust the factors to psychologically meaningful positions. Full statistical details of these operations will be found by the interested reader in Hildebrand (1953).

(ii) *Interpretation of the factors.* The four factors as finally stabilized accounted for 32% of the total variance. Factor loadings are given in Table 2. The first factor accounted for 12% of the total variance, and was easily identified as a neuroticism factor. As will be seen from Table 2, the four variables with the highest loadings differentiated significantly between normals and neurotics. Moreover, previous studies had shown these tests to be highly saturated in Neuroticism, emotional instability or similar factors.

The second factor extracted was identified as 'g' or general intelligence. This factor accounted for 9% of the total variance. The second factor was confidently identified as an intelligence factor. The first three variables were the three tests of intelligence included in the matrix, and a high negative correlation between intelligence and mistakes and length of time taken on the Porteus maze test seems appropriate.

The third factor extracted was identified as I-E. This factor accounted for 6% of the variance. For the purposes of this table critical ratios were calculated for the scores of the most highly loaded variables between the 45 hysterics and the 45 dysthymics. It will be seen that the tests do in fact discriminate between the groups. For these reasons the factor was identified as an I-E factor.

Table 2. *Final loadings for the four factors*

Variable	No.	Factors				h^2
		I	II	III	IV	
Prog. Matr.	1	179	750	-014	153	618
M.H.V.	2	-036	508	110	119	285
M.M.Q.	3	739*	-006	-257	055	615
Lie-scale	4	-380	050	-371	-009	284
Annoyances	5	429	-273	-012	068	263
S	7	739*	047	-472†	-043	773
T	8	467*	018	099	-142	248
D C	9	831*	-136	057	-127	728
R	10	-246	-236	563‡	-114	446
Power	11	098	744	012	-111	575
Sex humour	12	301	005	222	-198	179
Leg persistence	13	113	-007	-066	167	045
T.D. } Static	15	086	-082	180	520	317
R. } ataxia	16	-041	143	073	513	290
T.D. } Reversals	18	176	-006	052	292	118
R. }	19	021	161	261†	364	227
Body index	20	015	025	-255	035	067
Starting time	21	-167	055	-195	197	107
Total time	22	-062	-307	-071	321	206
Weighted wrong direction	23	257	-423	360†	-114	387
Lifted pencils	24	070	-211	006	098	059
Crossed lines	25	067	-168	425‡	-155	237
Totals		1·220	0·877	0·613	0·502	3·215

* The variables marked thus showed a significant difference $P > 0.01$ between normals and neurotics.

† Discriminates between the 45 hysterics and 45 dysthymics $P > 0.05$.

‡ Discriminates between the 45 hysterics and 45 dysthymics $P > 0.01$.

The fourth factor extracted was identified as arising from the intercorrelations of static ataxia and suggestibility scores. It does not seem possible to identify it more meaningfully.

(iii) *Regression coefficients and factor scores.* The factors having been identified, three deductions remained to be tested. It will be recalled that the experimental design called for discrimination between the various clinical groups by means of measurements made on the basis of scores based on the two factors of Neuroticism and I-E. It was felt that even though the factors were both significant and general they would have no more than a theoretical value unless such a discrimination could be made.

Ledermann's modification of Aitken's Pivotal Condensation method (Thomson, 1950) was used to estimate each individual's score on a given factor. The regression estimates so obtained 'are the best in the sense that they give the highest correlation, taken over a large number of men, between the estimates and the true value of a criterion when the latter can be separately ascertained'. As these requirements had been met in the present

experiment, the scores of all groups were converted to standard scores, the eight highest regression coefficients for the factors Neuroticism and I-E were selected, and the inner products of the standard scores and the regression coefficients calculated to give estimated factor scores for each member of the population on these two factors.

An overall analysis of variance was then made of the estimated scores on the Neuroticism factor, the results of which may be found in Table 3. As will be seen from Table 3, the analysis of variance showed no significant differences among the neurotic population, but a very significant difference when the normal group was included. A *t*-test was carried out between the neurotic population and the normal group and the c.r. equalled 3.74 which is highly significant. This is in correspondence with the experimental results and demonstrates that the battery does discriminate between normals and neurotics at a high degree of confidence.

Table 3

	Anx.	Obs.	Dep.	Hys.	P.P.	Mx.	Total	Normal
No.	25	10	10	25	20	55	145	25
Mean	5.949	5.424	4.647	4.866	5.916	5.364	5.409	3.402
S.D.	1.686	2.068	1.808	1.927	1.803	1.716	1.816	1.686

$F = 1.591$ with normals excluded, N.S.

$F = 5.858$ with normals included. Significant beyond the 1% level.

Table 4

	Anx.	Obs.	Dep.	Hys.	P.P.	Mx.	Total	Normal
No.	25	10	10	25	20	55	170	25
Mean	1.407	1.082	1.295	2.241	2.637	1.593	1.925	2.875
S.D.	0.986	1.489	1.131	1.271	1.384	1.384	1.399	1.084

$F = 5.896$, which is significant beyond the 1% level.

An analysis of variance was also carried out on the factor scores for the I-E factor. Results are found in Table 4. Inspection of the means demonstrates clearly that the distribution follows the prediction. The hysterics and psychopaths are clustered at the extravert end of the distribution, and the dysthymics at the introvert end of the distribution, with the mixed neurotics occupying a central position. The position of the normal group at the extravert end of the I-E continuum does not support 3. 3, but is consistent with the experimental results obtained from this, which suggested that as a whole they could be expected to be highly extraverted. *t*-Tests run between the groups showed no significant differences within the introvert and extravert clusters, but anxiety states were differentiated from hysterics at the 5% level and the other groups at higher levels of significance. Fig. 1 gives a graphic picture of the means of each group in relation to the two axes. The I-E axis has been drawn so as to cut off 10% of the normals and include them in the neurotic group, according to Russell Frazer's (1947) hypothesis that 10% of the normal population is in fact neurotic.

IV. DISCUSSION OF RESULTS

Jung's theory of Introversion-Extraversion is essentially descriptive in nature. The procedure used in the present study enabled the writer to support Jung's theoretical formulations on a more experimental basis than had hitherto been possible. It was found possible to divide the experimental population into three categories, which were in accor-

dance with the original hypothesis. These were neurotic introverts, neurotic extraverts, and the normal group.

(1) The introvert category is made up of anxiety states, reactive depressives and obsessionals. They score high on the Neuroticism factor and cluster at the I pole of the I-E factor. Descriptively Eysenck's amalgamation of the three groups into a single category of dysthymics seems substantiated by the experiment. On the factor scores there are no significant differences between the means of the groups making up this category. The factor-score criterion is probably a better one than one made on the basis of one or two tests.

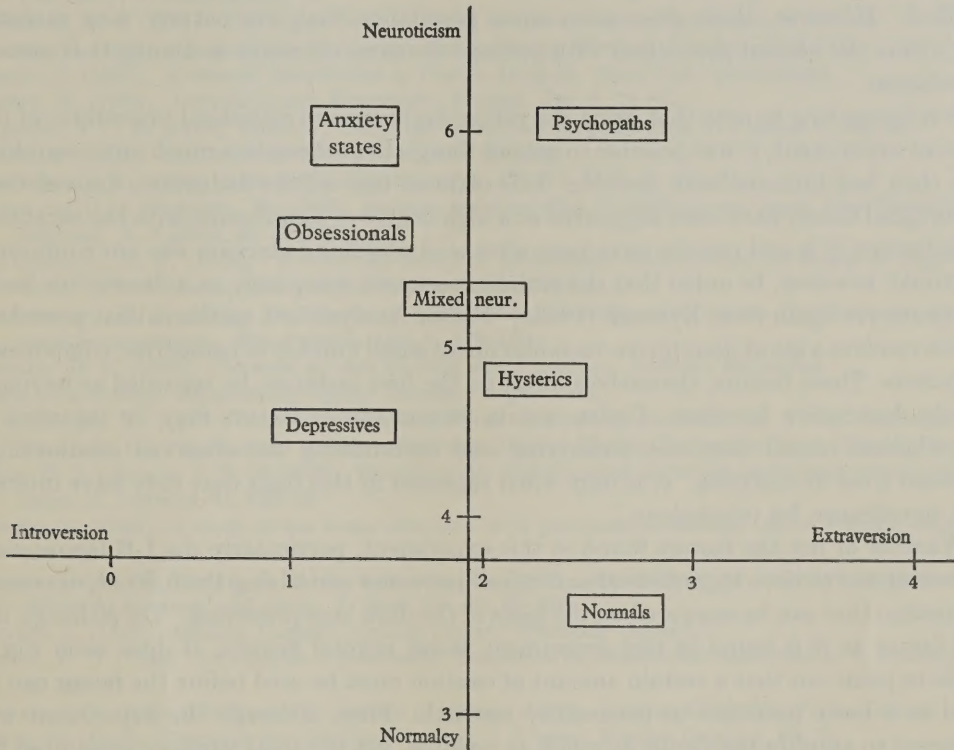


Fig. 1

(2) The extravert category is made up of hysterics and psychopaths. There is no significant difference between the means of the factor scores of these two groups. The actual result of the experiment which finds psychopaths scoring higher on the I-E factor than do the hysterics perhaps needs some discussion. It is inferred that the psychopaths were less inhibited in their test performance than were the hysterics, and it is interesting to note that Mayer-Gross, Moore & Slater (1949) found that on their second canonical variate, which seems to be a measure of I-E, the psychopaths scored higher than did the hysterics.

(3) The normal group did not in fact fulfil the prediction made for it originally. In terms of the other groups it was extremely extraverted although it differed significantly from the neurotic population on the Neuroticism factor. One possible explanation for this extreme degree of extraversion may perhaps be found in Eysenck's 1947 study of

ratings, where he found that such items as 'unskilled', 'narrow interests', and low I.Q. all had significant loadings on what he took to be an I-E factor. These three adjectives describe the control group well, and this may well have contributed to their degree of extraversion. Although it was not possible to test another group on the complete battery, another group of normals was tested on the R scale, which is the highest ranking test on the E pole of the I-E continuum. The control group was found to be significantly more extraverted than this second normal group (Hildebrand, 1956). It is possible that if a less homogeneous and more representative civilian sample of the normal population had been tested on the battery the predictions concerning this group would have been fulfilled. However, there does seem some possibility that the battery may measure I-E within the normal population with perhaps the same efficiency as among the neurotic population.

It is interesting to note that using the rather sophisticated statistical procedures of the present experiment, it was possible to submit Jung's hypothesis to a much more searching test than had hitherto been possible. It is claimed that all the deductions derived from the original theory have been supported at a high degree of significance, with the exception of deduction 3. 3, and reasons have been advanced why this deduction was not confirmed. It should, however, be noted that the entire experiment was purely on a descriptive basis. To quote yet again from Eysenck (1953c) 'Factor Analysis is a mathematical procedure which resolves a set of descriptive variables into a small number of categories, components or factors. These factors, themselves, must in the first instance, be regarded as having a purely descriptive function. Under certain circumstances factors may be regarded as hypothetical causal instances underlying and determining the observed relationships between a set of variables. It is only when regarded in this light that they have interest and significance for psychology.'

Whether or not the factors found in this experiment, particularly the I-E factor, may be considered to have hypothetical and causal instances underlying them is not, of course, a question that can be answered on the basis of the data here presented. Yet although the I-E factor as it is found in this experiment seems beyond dispute, it does seem worth while to point out that a certain amount of caution must be used before the factor can be used as a basic postulate in personality research. First, although the experiment was designed to amplify the factor as much as possible, yet the total variance accounted for by the factor amounted to only 6%, only 1% more than that of a factor that it was found impossible to identify satisfactorily. Secondly, of the tests which were suggested as possible measures of I-E, only two differentiated between introverts and extraverts at the 1% level of significance. That, in spite of these results, the battery was as successful as it was is very satisfactory, but the next step in any further investigation of this area of personality would seem to be further descriptive studies rather than a search for the hypothetical causative instance.

This article is based on the writer's unpublished Ph.D. thesis for London University. The writer wishes to thank Dr L. Minski, Physician Superintendent of Belmont Hospital, for kind permission to test patients within the hospital, and Dr M. M. Desai and Mr R. Secker of the Psychology Department, Belmont Hospital, for their many kindnesses. He also wishes to thank his supervising teacher, Dr H. J. Eysenck, for his advice and assistance.

REFERENCES

- BARTLETT, M. S. (1950). Tests of significance in factor analysis. *Brit. J. Psychol. Statist. Sect.* **3**, 77-85.
- BURT, C. (1952). Tests of significance in factor analysis. *Brit. J. Psychol. Statist. Sect.* **5**, 109-33.
- CATTELL, R. B. (1946). *The Description and Measurement of Personality*. New World Book Co.
- COLLIER, R. & EMCH, M. (1938). Introversion-Extraversion—the concepts and their clinical use. *Amer. J. Psychiat.* **94**, 1045-75.
- EYSENCK, H. J. (1947). *Dimensions of Personality*. London: Kegan Paul.
- EYSENCK, H. J. (1953*a*). *The Structure of Human Personality*. London: Methuen.
- EYSENCK, H. J. (1953*b*). Research Report under contract No. 62558s-56 with the Bureau of Medicine, U.S. Navy. (Unpublished.)
- EYSENCK, H. J. (1953*c*). The logical basis of factor analysis. *Amer. Psychologist*, **8**, 105-14.
- FRAZER, R. (1947). *The Incidence of Neurosis among Factory Workers*. London: H.M.S.O.
- FREUD, S. (1920). *A General Introduction to Psycho-Analysis*. New York: Liverwright.
- FREYD, M. (1924). Introverts and Extroverts. *Psychol. Rev.* **5**, 74-87.
- GARRETT, J. C. M. (1918). General Ability, cleverness and purpose. *Brit. J. Psychol.* **9**, 345-66.
- GROSS, O. (1902). *Die Cerebrale Sekundarfunktion*. Leipzig.
- GUILFORD, J. P. (1934). Introversion-Extraversion. *Psychol. Bull.* **31**, 331-54.
- HEYMANS, G. & WIERSMA, E. (1906). Beitrage zur speziellen Psychologie auf grund einer Massenufer-suchung. *Z. Psychol.* **42**, 81-127 *et seq.*
- HILDEBRAND, H. P. (1953). 'A Factorial Study of Introversion-Extraversion by Means of Objective Tests.' Ph.D. Thesis, University of London. (Unpublished.)
- HILDEBRAND, H. P. (1956). On the use of Guilford's introversion-extraversion Questionnaire with a neurotic population. *Rev. Psychol. App.* **6**, 179-185.
- JORDAN, W. F. (1890). *Character as seen in Body and Parentage*. London: Macmillan.
- JUNG, C. J. (1924). *Psychological Types*. London: Kegan Paul.
- LAWLEY, D. N. (1950). *The Maximum Likelihood Method of Estimation of Factor Loadings*. In Thomson, G.: *The Factorial Analysis of Human Ability*. London University Press. (5th ed.)
- LINE, W. & GRIFFIN, J. D. M. (1935). The objective determination of factors underlying mental health. *Amer. J. Psychiat.* **91**, 833-42.
- LOVELL, C. (1945). A study of the factor structure of 13 personality variables. *Educ. Psychol. Meas.* **5**, 335-50.
- MAYER-GROSS, W., MOORE, J. T. N. & SLATER, P. (1949). Forecasting the incidence of neurosis in officers in the army and navy. *J. Ment. Sci.* **95**, 80-100.
- MINSKI, L. (1950). *Belmont Hospital, Annual Report*. (Privately printed.)
- NORTH, R. D. (1948). An analysis of personality dimensions of introversion-extraversion. *J. Personality*, **17**, 552-67.
- THOMSON, G. (1950). *The Factorial Analysis of Human Ability*. London University Press. (5th ed.)
- WEBB, E. (1915). Character and intelligence. *Brit. J. Psychol. Monogr. Suppl.* **1**, No. 3.

(Manuscript received 4 October 1956)

DETERMINANTS OF EMOTIONALITY IN THE RAT

II. ANTECEDENT FACTORS*

By P. L. BROADHURST

University of London, Institute of Psychiatry (Maudsley Hospital)

A total of 112 rats (56 males and 56 females) is used to investigate the effects of some antecedent factors upon the defecation and ambulation scores derived from Hall's open-field test of emotionality, given under precisely standardized conditions. Conditions of rearing and husbandry are similarly constant, except where systematically varied for experimental purposes. The results show that neither defecation nor ambulation is affected by systematic 'gentling'. Sexual experience does not affect defecation, but increases ambulation significantly, and this effect appears due, among females, to the experience of suckling a litter. Previous experience of the same test at earlier ages decreases defecation among young rats, but not among older ones, while ambulation is unaffected throughout. The effect of ageing itself upon defecation is seen as a decrease among older rats only, young ones being unaffected: conversely, its effect upon ambulation is an increase among young ones only. The previously reported sex difference in response is confirmed; no interaction between this sex difference and any of the experimental variables was detected. It is concluded that these findings further support the status of defecation, as observed in this test, as a measure of an emotional response.

I. INTRODUCTION

In a previous paper (Broadhurst, 1957*a*), it was shown that the defecation score in Hall's open-field test (1934) provides a valid measure of emotional response in the rat, and the effect upon it of several variables within the test situation was measured. For this purpose, the previous experience and husbandry of the subjects was carefully kept uniform. In the present paper, the effect of some variables antecedent to the test situation is examined. It is not claimed that this survey of the non-situational determinants of emotionality will be in any way exhaustive; the choice of those reported here has been determined principally by practical considerations relating to further use of the test.

II. APPARATUS AND PROCEDURE

Because it is the independent variables outside the test situation which are to be manipulated in the present investigation, complete uniformity of the open-field test conditions is necessary. Accordingly, the standardization of the test conditions developed in the previous paper (Broadhurst, 1957*a*) was employed throughout, and is briefly specified below. Above a circular arena, 32 $\frac{3}{4}$ in. in diameter with white plywood walls 12 $\frac{1}{2}$ in. high, a battery of loudspeakers and of photographic lamps provided sound ('white' noise) and light fields whose intensity at floor-level averaged 78 db (ref. 0.0002 dynes/sq. cm.) and 165 cp. respectively. (This is the low level of noise used in the first experiment in this series; see Broadhurst 1957*a*, p. 3. This revised determination was made possible by more accurate apparatus which, however, confirms the figure of 94 db previously given for the high level but not now used.) The subjects were transported individually in a covered white enamel pail to the test arena where they remained for 2 min. a day on four successive days. The number of fecal boluses deposited and the number of compartments marked on the floor and roughly equal in size, through which the animal passed, constitute the basis for the defecation and ambulation scores respectively.

Apart from experimental variations in the pre-test experience of the subjects which are the object of investigation and which are described in the appropriate section below, the standard conditions of husbandry used in the previous study (Broadhurst, 1957*a*) obtained. Salient features are the minimum possible handling, together with uniformity of diet and of caging after weaning at 21 days of age until testing took place in the open field as soon after 100 days of age as was practicable.

* This paper is based on part of a thesis approved by the University of London for the degree of Ph.D.

(i) Introduction

In the present section the effects of three antecedent variables are considered. The first of these is overt sexual experience. Sexual deprivation, it is thought, may lead to subsequent neurosis (e.g. 'The neurosis owes its existence to a diversion and regression of the libido from the genital position...' (Reich, 1950, p. 131)). While it is not claimed either that sexuality plays as important a part in the experience of the rat as it tends to do in a behaviourally more complex organism like man, or that emotionality as measured by the open-field test can be equated directly with human personality dimensions, it is clear that segregated rats are deprived of a social experience unique in their repertoire of behaviour. Furthermore parturition and/or suckling in females may affect the metabolic balance (Watson, 1905) and thereby alter emotional elimination which is under the control of the sacral branch of the sympathetic nervous system (Hall, 1934).

Practically, it was desirable to know whether or not using our subjects for breeding before the open-field test at the standard 100 days of age would affect their performance thereon. The fact that the social stimulus of another animal in the test situation itself has no effect on scores (Anderson, 1939) suggests that non-situational social experience may be unimportant. The further possibility exists of investigating the effects of pregnancy and parturition among females independently from the effects of suckling a litter, by depriving some subjects of their litters at birth and allowing others to keep them until the usual weaning time, and among the males of investigating any differences between the effect of allowing mating with females alone or in sexual competition with other males.

The second antecedent variable is handling or 'gentling', as it is more expressively termed in American laboratories. It is a common pre-experimental routine in many animal studies, for it renders the subjects more docile, less likely to bite, and reduces flight or startle reactions, including emotional elimination, when touched. The extinction of such behaviour by systematic handling, stroking and petting is therefore of practical convenience in the carrying out of experimentation and is widely employed. However, as Christie (1951) has emphasized, the possible effect of such training on experimental results is frequently ignored. Evidence suggests that it may be important (Bernstein, 1952, 1956; Scott, 1955; Ruegamer, Bernstein & Benjamin, 1954; Levine, Chevalier & Korchin, 1956; Stanley & Monkman, 1956).

With respect to emotionality, Weininger (1954) has shown that gentled rats are significantly more active in the open field on the first day of testing than a group not so treated. There is also a significantly greater weight gain in gentled animals (Weininger, 1954; Bovard, 1954; Weininger, McClelland & Arima, 1954). On the other hand, Hunt & Otis (1955) failed to find differences in defecation scores attributable to gentling, and Gertz's findings (personal communication, 1956) suggest that the effect of gentling may be temporary.

Related to the effect of gentling is that of the third experimental variable to be considered, namely, previous experimental experience. The only systematic investigation appears to be that of Watson (1954*a*, 1954*b*) whose unpublished data show that animals who, after maturity, have undergone training in other experimental situations or have been given the open-field test itself display significantly lowered emotional reactivity when subsequently tested. This demonstration of 'experimentation sophistication', as Watson terms it, is clearly of great practical importance in planning animal studies

involving sequential testing and in interpreting their results. It is desirable, therefore, to see if the effect holds with female subjects and with experience of testing in early life.

(ii) *Experimental design*

The experimental design is shown in Table 1, from which it will be seen that of the three independent variables whose effect on the emotional reactivity scores of rats of both sexes it is desired to establish, that is, sexual experience, handling, and previous test experience, the last two were combined as aspects of one variable, since these two treatments seemed most alike, and hence any possible interactions between them would be of the least interest. This was done because there were insufficient subjects available to allow six replications of a complete fourfold classification in an analysis of variance design.

Table 1. *Experimental design showing treatments imposed before testing*

	Sexual experience		No sexual experience	
	Males*	Females	Males	Females
Handling				
Previous test experience	Males	Females	Males	Females
Neither above, i.e. a control group with no handling and no previous experience	Males	Females	Males	Females

* The *n* for each sub-group is 6, total *N* = 72.

(iii) *Subjects and procedure*

Since a decrease in defecation scores was anticipated from the experimental treatments, animals from a high scoring or emotionally reactive strain being developed by selective breeding (Broadhurst, 1957c) were used in so far as they were available. The 72 subjects for this experiment thus constituted part of the first generation of the selective breeding study of emotionality. The 60 animals comprising the experimental groups with pre-test treatments were drawn from a number of animals superfluous to that study and control groups without treatment were provided by randomly selecting 12 animals (6 males and 6 females) from among the main populations of animals routinely tested for the purpose of the breeding study. Appropriate *t*-tests yield non-significant values showing that these control groups impartially represent the main populations in respect of defecation and ambulation scores. As a further check, the possibility of a hereditary bias influencing the scores of the experimental groups was investigated by comparing the scores obtained by the parents of these animals with those of the subjects retained in the selective breeding experiment. In all important respects no significant differences were found. Further details of these checks on the justification for regarding the subjects used as representative of the larger populations from which they came are given elsewhere (Broadhurst, 1956).

Variations from the normal procedures of rearing the animals constitute the experimental treatments and were as follows. (1) *Handling* or gentling followed Weininger's (1954) procedure. Each rat was taken individually from its cage each day and held in the experimenter's lap with his left hand while being stroked from head to tail with his right thumb for 5 min. at a rate of about 90 strokes per minute. This procedure was begun when the 24 animals in this group were 56.5 days old (± 0.16) and continued daily for 57 days until they were given the open-field test. (2) *Previous experience of the test* was ensured by giving the open-field test to the 24 animals in this group, under standard conditions and beginning when they were 53.3 days old (± 0.22). (3) *Opportunity for sexual experience* was allowed simultaneously to all the 36 subjects constituting this group when they were 64 days old (± 0.13). One male was randomly selected from the normal cage complement of three and was mated with 2 females in their home cage. His place was taken by the single female thus displaced from the second cage. Four days later the original caging was restored. This procedure was carried out three times beginning at intervals of 1 week. Thus, the single males (a total group of 6 rats) were always mated with 2 females each, their two cage mates (a total group of 12) were always mated with 1 female per pair. The females concerned, however, were rotated to form these sub-groups. As a result of these matings, all females, except one, became pregnant within 10 days of first mating, and were removed to individual cages before their litters were born. Half of them (9 rats) had their litters augmented to a standard size of 12 pups within 3 days of parturition,

and were permitted to rear them to the normal age of weaning of 21 days, which they all did successfully. The rest of the females constituted the suckling-deprived group, and either were deprived of their pups (7 cases) or lost them (1 case) within 3 days of parturition, or did not become pregnant (1 case).

All animals were tested in the open field under the standard conditions described in section II above when their mean age was 111.4 days (± 1.71). There was no significant difference between males and females in this respect.

(iv) Results

The mean scores for the various sub-groups, together with a summary of the results of the analyses of variance performed on the two scores, are shown in Table 2. The interactions (three first order and one second order) were uniformly non-significant and are

Table 2. *Mean scores of sub-groups and summary of results of analyses of variance*

		Score				
		Defecation (av. no. of fecal boluses per day)		Ambulation (av. no. of metres per day)		
Source of variance (treatments)	Sub-group	Mean	Significance of <i>F</i> ratio	Mean	Significance of <i>F</i> ratio	D.F.
Sexual experience	Yes	3.2	N.S.	8.7	<i>P</i> < 0.01	1,60
	No	2.6		6.9		
Other treatments	Handling	3.3	<i>P</i> < 0.01	8.8	N.S.	2,60
	Previous testing	2.1		7.0		
	Neither	3.3		7.6		
Sex	Males	2.8	N.S.	7.1	<i>P</i> < 0.05	1,60
	Females	3.0		8.5		

Table 3. *Significance of differences in defecation scores between different treatments*

Comparison	Mean difference	Significance
Previous testing <i>v.</i> Handling	1.2	$P < 0.01$
Previous testing <i>v.</i> Neither treatment	1.1	$P < 0.01$
Handling <i>v.</i> Neither treatment	0.1	N.S.

therefore omitted from the table. The significance of the *F*-ratio associated with the main effect upon defecation scores designated 'other treatments' permits the application of *t*-tests to evaluate the significance of the differences (one-tail tests of significance were used where appropriate) between the defecation scores of the three sub-groups constituting this variable. The results are given in Table 3. In a further subsidiary analysis of the group allowed sexual experience *t*-tests were used to investigate the differences between the two groups given slightly different sexual experience, since these sub-groups could not have been conveniently differentiated in the main analysis (Table 2). The outcome is shown in Table 4.

These results clearly fail to confirm the reported decrease in emotionality following gentling, despite its effect upon growth. The effect upon weight—both males and females at testing were heavier than ungentled animals—was not, however, significant. The only experimental variable which has had an effect upon the defecation scores is that of previous experience of the test. Animals which have been given the test once before at the age of some 53 days defecate less than those undergoing it for the first time. (The effect of re-testing later in life will be discussed in the next section.) The usual sex difference in defecation scores is not seen, probably because the emotionally reactive strain, from

Table 4. *Significance of differences between groups given slightly different sexual experience*

Sex	Treatment	n	Score			
			Defecation (av. no. of fecal boluses per day)		Ambulation (av. no. of metres per day)	
			Mean	Significance of difference	Mean	Significance of difference
Males	Mated in pairs with single female	12	3.0	N.S.	8.4	N.S.
	Mated alone with two females	6	2.6		7.7	
Females	Allowed to suckle litter	9	3.4	N.S.	10.4	$P < 0.01$
	Deprived of litter	9	3.6		8.1	

which most—54 out of 72—of the subjects were drawn, does not show this difference as clearly as does the non-emotional strain, or the unselected populations used in the previous study (see Broadhurst, 1957*c* and 1957*a*, respectively). This interaction between strain and sex differences in selectively bred animals is fully discussed elsewhere (Broadhurst, 1958): there seems no reason to suppose it has any bearing on the outcome of the present experiments.

The ambulation scores, which do not in general display this interaction between sex and strain, show the usual sex differences, the females being significantly more active than the males. In contrast with its negligible effect upon defecation scores, the sexual experience variable has an important effect upon ambulation. With respect to the females' experience, it is possible to be precise about which of the inter-group differences is associated with this effect, since the females without sexual experience have a mean ambulation score of 7.8 m. per day, which is insignificantly different from that of the litter-deprived group of females (Table 4). This suggests that it is the experience of suckling a litter which is crucial.

III. EFFECTS OF AGE

(i) *Introduction*

This section is concerned primarily with the effect of age upon open-field test scores. Small variations in age of testing mature animals have no effect (Broadhurst, 1957*a*), but O'Kelly's results (1940) suggest that young rats show a decrease in defecation scores with age of testing during early life. The most extensive analysis of the effects of age on scores of mature (100 days and older) rats is provided by Watson's unpublished data (1954*b*). For example, his results from four groups of males showed a systematic increase of defecation scores with advancing age from 131 to 341 days, and the difference between the extreme groups reaches the 3% level of significance. It is proposed to attempt to extend these findings, first, by including both immature animals and mature ones of an age intermediate between Watson's extreme groups, and secondly, by using both sexes in the experimental analysis. In view of the highly significant findings relating to previous test experience among young rats and reported in the previous section, the opportunity will be taken to examine the effect of such experience upon testing much later in life.

In effect, these objectives were accomplished by dividing the investigation into two parts, and investigating first the effects of ageing before the standard age of testing had

been reached, and secondly, the effects of ageing after maturity. Since the previous work would lead one to predict a decrease of emotional defecation with age in young rats and the reverse in mature ones, animals from the high-scoring, or emotionally reactive strain, derived from the selective breeding experiment mentioned earlier, were used for the first part, and animals from the low-scoring, or non-reactive strain for the second. In this way the expected effects could be most clearly shown.

(ii) *Subjects and procedure*

All subjects were reared and tested under the standard conditions described earlier (section II). The only variations—which constitute the independent variables to be studied in this section—lay in the testing, or in some cases re-testing, at ages systematically different from the standard one of around 100 days.

For the first part of the experiment, the subjects were the 12 male and 12 female rats of the emotionally reactive strain who constituted the 'previous test-experience group' in the previous section. For that purpose they were tested at a mean age of 53.3 days before being re-tested 60 days later at a mean age of 113.3 days (± 0.22).

The subjects for the second part of this experiment were 10 male and 10 female rats of the second generation of the emotionally non-reactive strain whose open-field testing was delayed until a mean age of 242.3 days (± 0.17). For comparison, control groups also of 10 male and 10 female non-emotional rats were selected. Their mean defecation and ambulation scores did not differ significantly (by *t*-test) from those of the main population, also tested at the standard age. These control groups were given the test for the second time at a mean age of 237.6 days (± 0.60).

(iii) *Results*

The outcome of the comparisons relating to the younger animals is presented in Table 5. Because the handling variable is the only one in which the group tested at 50 days of age differs from all the rest of the subjects used in the experiment reported in the previous section—numbers of each sex, and of those allowed sexual experience being proportion-

Table 5. *The effects of ageing on the test scores of young rats*

Score	Approximate age at 1st testing		Significance of difference, 50 v. 100 days	Approximate age at 2nd testing (100 days)	Significance of difference, 50 v. 100 days	Significance of difference, 1st v. 2nd testing at 100 days
	50 days	100 days				
Defecation (av. no. of fecal boluses per day)	3.1	3.3	N.S.	2.1	$P < 0.01$	$P < 0.01$
Ambulation (av. no. of metres per day)	6.0	8.3	$P < 0.001$	7.0	$P < 0.001$	N.S.

ately equal—and because handling has been shown to have had no effect, these latter subjects therefore form an appropriate control group for the results of testing at the standard age (see col. 3). The differences for each score, evaluated by *t*-test, are shown in the fourth column. The rest of the table shows the effect of re-testing, that is, a comparison is made of the scores returned at 50 and 100 days of age by the same subjects, these differences being accordingly evaluated by Sandler's (1955) *A*-test for correlated measures (see col. 6). They should be compared with those between naïve and test-sophisticated rats of the *same* age derived from the analyses of the previous section (Tables 2 and 3) and which are reproduced for convenience in the final column.

The results of the comparisons relating to the older age groups are presented in similar form in Table 6. The significance of the difference between the scores of the same animals

on the two occasions of testing was evaluated as before with the *A*-test (col. 6); otherwise double classification analyses of variance were used in which the sex difference was included as the second variable. The significance of the *F*-ratios associated with the main effects due to age of testing or the previous test experience—the other variable in the analysis of variance—are indicated in the table (cols. 4 and 7 respectively), but, as might have been expected from these subjects belonging to the emotionally non-reactive strain (Broadhurst, 1958), the sex differences were significant throughout and in the usual direction—males defecating more than females and ambulating less—and are not therefore shown in the table. There were no significant interactions between age and sex.

Table 6. *The effects of ageing upon the test scores of older rats*

Score	Approximate age at 1st testing		Significance of difference, 100 v. 240 days	Approximate age at 2nd testing (240 days)	Significance of difference, 100 v. 240 days	Significance of difference, 1st v. 2nd testing at 240 days
	100 days	240 days				
Defecation (av. no. of fecal boluses per day)	2.6	1.5	$P < 0.05$	1.3	$P < 0.001$	N.S.
Ambulation (av. no. of metres per day)	7.8	7.6	N.S.	7.4	N.S.	N.S.

Trends in open-field test scores with age and test experience can be discerned in these results, and the comparison of the first test and the re-test at the different ages permits a rough analysis of the influence attributable to these two factors, which may be stated as follows. First, it is clear that there is little change in defecation scores between 50 and 100 days, whereas between 100 and 240 a significant decline occurs. But, while testing at 50 days causes a significant decrease on re-testing at 100 days, testing at the latter age has no additional influence on the decline in defecation scores at 240 days. Second, though ambulation scores increase significantly between 50 and 100 days of age, they change little between 100 and 240 days, and the effect of prior testing at either 50 or 100 days upon later performance is negligible. There is no evidence suggestive of sex differences in any of these respects, but it must be stressed that these findings may have limited generality because animals with characteristically high or low defecation scores were used in the first and second parts of this study respectively, and it has not been established that different strains of animals do not vary differently with age. Indeed the fact that the present findings run counter to those of both O'Kelly (1940), and Watson (1954*b*) suggests that such may be the case, though differences in testing technique may also be important.

V. DISCUSSION AND CONCLUSIONS

The findings from the experiments reported in this paper suggest the following conclusions: the only antecedent variables shown to be indubitably related to emotional reactivity as measured by defecation scores are prior experience of the same test, and age after maturity. With respect to others investigated, it is, of course, only possible to say that their effect, if any, has not been detected under what seem to be favourable conditions. Thus, further work, reported elsewhere (Broadhurst, 1956), shows, first, that depriving rats of form vision by screening the front of their cages with opaque Perspex throughout their life before testing has no effect on open-field test scores, and secondly,

that such scores are not related to rodent body-build, as measured by Watson's (1957) index.

The confirmation of the test-experience effect described by Watson (1954*a*), though partial—it was found to be present among young rats but absent in ones of a similar age to his subjects—strengthens the conclusion that previous test experience does lower defecation scores on subsequent testing. The explanation is not far to seek: adaptation to the mild stress situation represented by the open-field test is clearly demonstrated by the well-established decrease in defecation scores with repeated testing on successive days—indeed it is this extinction which supports the validity of defecation as a fear response. Clearly it may be anticipated that part of such adaptive learning derived from one series of daily tests will be carried over to a subsequent series, thus giving a reduced score. The fact that learning situations different from the open-field test have a similar effect (Watson, 1954*b*) may be expected from the generality of the defecation response to strange situations (Parker, 1939).

But it is this generality which would lead one to expect that animals, gentled so that they—*inter alia*—no longer defecate when handled, would not do so when faced with the open-field situation, whereas our findings fail to confirm those of others which suggest the effect mentioned. However, gentling in most of these cases was started at an earlier age than that used here, and it may well be that 'critical periods' of susceptibility to certain kinds of learning occur in the life of the rat, just as they have been shown to do in other species (Scott & Marston, 1950; Tinbergen, 1951, p. 150). It is clear that some such hypothesis relating to the effects of handling upon the increase in body-weight in rats would reconcile the finding that 21 days of gentling begun at age 23 days increases adult body-weight significantly (Weininger, 1954), whereas 57 days of gentling begun at age 56 days did not. Similarly, the result that, among our subjects, prior test experience has a potent effect if given at 50 days, but not if given at 100 days of age accords with this view.

It has thus been shown that the defecation scores derived from the standardized open-field test used here are insensitive to variations in several kinds of antecedent factors which might have been expected to affect them. These findings consequently further enhance the usefulness of the measure as an index of an emotional response. Only previous experience of the test and large differences in age after maturity have been shown to be important in this respect, and the effect of these variables, like that of the sex difference previously reported (Broadhurst, 1957*a, b*) and, in general, confirmed here, must be taken into account in making practical use of the test.

The ambulation scores represent a less important index of emotionality, since they largely reflect the intensity of evocation of exploratory tendencies (Broadhurst, 1957*a*). Consequently they are of less interest in the present connexion. The findings differ, however, from those relating to defecation in two principal ways. First, there is no evidence that ambulation is influenced at all by previous testing. This suggests that the exploratory drive is easily aroused but is weak, since learning based upon its reduction is impermanent. Secondly, there is a marked effect attributable to previous experience of a sexual nature. This can be interpreted in several ways. The absence of corresponding changes in defecation scores argues against its being due to a decrease in fearfulness, while the specificity of the causative factor attributable in females—the experience of suckling a litter—suggests a change in activity level, perhaps learned, or perhaps metabolically

determined. Why the experience of mating should similarly affect males is not apparent from these data: what is clear is that further work would be needed to confirm and elucidate these particular findings.

REFERENCES

- ANDERSON, E. E. (1939). The effect of the presence of a second animal upon emotional behavior in the male albino rat. *J. Soc. Psychol.* **10**, 265-8.
- BERNSTEIN, L. (1952). A note on Christie's: 'Experimental naïveté and experiential naïveté'. *Psychol. Bull.* **49**, 38-40.
- BERNSTEIN, L. (1956). A reply to Scott. *J. Abnorm. Soc. Psychol.* **53**, 141-2.
- BOVARD, E. W. (1954). A theory to account for the effects of early handling on viability of the albino rat. *Science*, **120**, 187.
- BROADHURST, P. L. (1956). *Emotionality in the Rat: its Determinants, Inheritance, and Relation to some Aspects of Motivation*. Ph.D. thesis, University of London.
- BROADHURST, P. L. (1957*a*). Determinants of emotionality in the rat: I. Situational factors. *Brit. J. Psychol.* **48**, 1-12.
- BROADHURST, P. L. (1957*b*). Determinants of emotionality in the rat: III. Strain differences. *J. Comp. Physiol. Psychol.* (In the Press.)
- BROADHURST, P. L. (1958). Emotionality in the domestic rat as modified by selective breeding. (In the Press.)
- CHRISTIE, R. (1951). Experimental naïveté and experiential naïveté. *Psychol. Bull.* **48**, 327-39.
- HALL, C. S. (1934). Emotional behavior in the rat. I. Defecation and urination as measures of individual differences in emotionality. *J. Comp. Psychol.* **18**, 385-403.
- HUNT, H. F. & OTIS, L. S. (1955). Restricted experience and 'timidity' in the rat. *Amer. Psychologist*, **10**, 432. (Abstract.)
- LEVINE, S., CHEVALIER, J. A. & KORCHIN, S. J. (1956). The effect of early shock and handling on later avoidance learning. *J. Personality*, **24**, 475-93.
- O'KELLY, L. I. (1940). The validity of defecation as a measure of emotionality in the rat. *J. Gen. Psychol.* **23**, 75-87.
- PARKER, M. M. (1939). The interrelationship of six different situations in the measurement of emotionality in the adult albino rat. *Psychol. Bull.* **36**, 564-5. (Abstract.)
- REICH, W. (1950). *Character Analysis* (3rd ed.). London: Vision Press.
- RUEGAMER, W. R., BERNSTEIN, L. & BENJAMIN, J. D. (1954). Growth, food utilization, and thyroid activity in the albino rat as a function of extra handling. *Science*, **120**, 184-5.
- SANDLER, J. (1955). A test of the significance of the difference between the means of correlated measures, based on a simplification of Student's *t*. *Brit. J. Psychol.* **46**, 225-6.
- SCOTT, J. H. (1955). Some effects at maturity of gentling, ignoring, or shocking rats during infancy. *J. Abnorm. Soc. Psychol.* **51**, 412-14.
- SCOTT, J. P. & MARSTON, M. (1950). Critical periods affecting the development of normal and mal-adjustive social behavior of puppies. *J. Genet. Psychol.* **77**, 25-60.
- STANLEY, W. C. & MONKMAN, J. A. (1956). A test for specific and general behavioral effect of infantile stimulation with shock in the mouse. *J. Abnorm. Soc. Psychol.* **53**, 19-22.
- TINBERGEN, N. (1951). *The Study of Instinct*. Oxford: Clarendon Press.
- WATSON, J. B. (1905). The effect of the bearing of young upon the body-weight and weight of the central nervous system of the female white rat. *J. Comp. Neurol. Psychol.* **15**, 514-24.
- WATSON, R. H. J. (1954*a*). The effects of age and experience on the performance of the white rat on a test of emotionality. *Bull. Brit. Psychol. Soc.* **23** (inset), 6. (Abstract.)
- WATSON, R. H. J. (1954*b*). *The Behaviour of the Rat in Conflict Situations Presented Serially*. Ph.D. thesis, University of London.
- WATSON, R. H. J. (1957). A factor analysis of body measurements in the rat. (In preparation.)
- WEININGER, O. (1954). Physiological damage under emotional stress as a function of early experience. *Science*, **119**, 285-6.
- WEININGER, O., MCCLELLAND, W. J. & ARIMA, R. K. (1954). Gentling and weight gain in the albino rat. *Canad. J. Psychol.* **8**, 147-51.

(Manuscript received 12 November 1956)

SURPRISE, ISOLATION, AND STRUCTURAL CHANGE AS FACTORS AFFECTING RECALL OF A TEMPORAL SERIES

BY ROBERT T. GREEN

Psychology Department, University College, London

SUMMARY

Two experiments are considered which throw light on certain aspects of what has become known as the von Restorff effect. The first is designed to find out whether items in a serial learning task are favoured in recall because unexpected change attracts attention. The results show that change as such, expected or unexpected, has an effect, though unexpected change does have a somewhat greater effect on recall scores. Subjective reports suggest that this is due, at least in part, to the fact that when the instructions are understood they serve to direct attention. The second experiment directly attacks the question of whether 'isolation' has any distinct effect on probability of recall, and demonstrates that structural change in a temporal series is the crucial variable.

EXPERIMENT 1. THE EFFECT OF INSTRUCTIONS INTENDED TO REDUCE THE 'SURPRISE' VALUE OF STRUCTURAL CHANGE

I. *Introduction*

In a previous paper (Green, 1956) evidence was considered which indicated that an unexpected change in structure of a temporal series increased the probability of recalling the item introducing the change. The degree of 'surprise' occasioned by a particular change was inferred by assuming that subjects' expectations were formed by an inductive process. Thus a particular nonsense syllable occurring in the fourth serial position was more likely to be recalled if it were preceded by three items of another sort (in this case three-digit numbers) than if it were preceded by three other nonsense syllables. Later changes of this kind in the same list would be less unexpected since once the uniform series is broken the inductive principle would lead to a recognition of the possibility of similar changes. Put more simply: after the first change the subject would know he was dealing with a mixed list. The data seemed to support this contention inasmuch as the first change showed a large differential recall effect whereas later changes produced progressively smaller differences in recall scores.

In order to verify this 'surprise' hypothesis more directly it was decided to run two more groups of subjects under similar conditions to those holding for half the subjects in the original experiment, with the exception that the 46 new subjects were given detailed information about the structure of the list immediately prior to presentation. It was predicted that, if surprise is an important factor and if the information reduces or eliminates surprise, then the item introducing the change would be recalled less often by the subjects with knowledge of structure than those subjects without such knowledge. In fact it should be recalled no better than by the groups for whom the item introduced no change.

II. *Experimental procedure*

The experiment consisted in displaying 20 items consecutively for 5 sec. each and requesting the subjects to recall in writing as many of the items as possible without regard to order. The procedure was identical with that described in the original paper except that in the instructions given to 23 of the subjects the

following addition was made. 'Finally, let me say something about the way the list is made up so you will have some idea of what to expect. In all there are 20 items, one on each card. I shall expose these cards one at a time for a period of 5 sec. each. On half of the cards there is a nonsense syllable. The other ten carry three-digit numbers. A nonsense syllable is a vowel with a consonant either side such as *BOK*. This particular nonsense syllable will not appear in the list, but 10 of the items will be of a similar nature. An example of a three-digit number would be *395* which again will not appear, although the other 10 items will be of this sort. The items are presented in a particular order. The first three will be nonsense syllables, then a number, followed by six more nonsense syllables, then six numbers, another nonsense syllable, and a final three numbers.' A similar addition was made for the other 23 subjects except that 'nonsense syllable(s)' became 'number(s)' and vice-versa.

So the 46 subjects given the lists *a* or *A* in the original experiment were matched by two groups of the same sizes on the same lists. For the sake of convenience, the lists *a* and *A* are shown again in Table 1. As may be seen, the crucial item in position 4, which for group *A* is *406* and for group *a* is *HOF*, is surrounded by items of the other sort. The sample consisted mainly of university graduates and students with the addition of a few experimentally naïve adults from other sources to round off the number in each group.

Table 1. *Structure of the two lists*

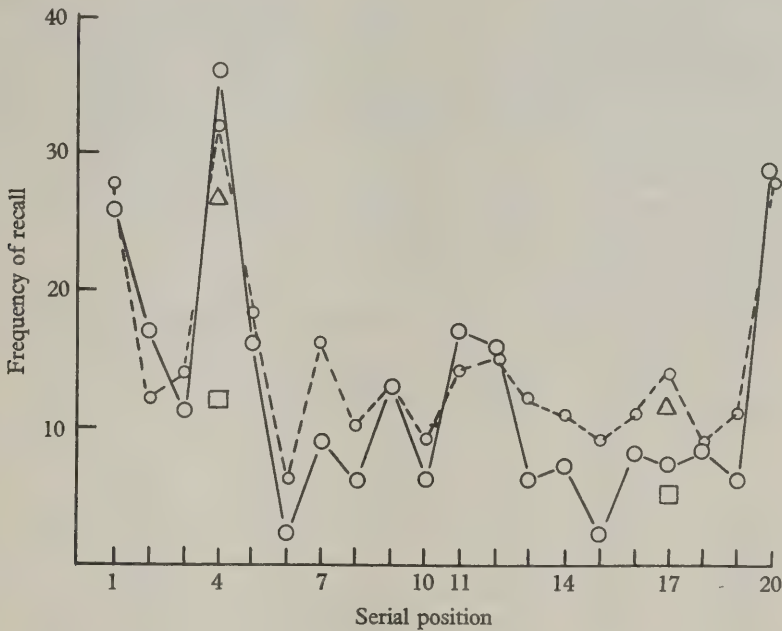
		List A	
	↓		
	1	GUB	20
	2	KEV	19
	3	DAC	18
	4	406	17
	5	TER	16
	6	WAJ	15
	7	SIH	14
	8	RUL	13
	9	VOM	12
	10	FIP	11
	11	417	10
	12	269	9
	13	341	8
	14	562	7
	15	738	6
	16	179	5
	17	HOF	4
	18	258	3
	19	763	2
	20	581	1
			↑
		List a	

Interest in the results was mainly centred on the probability of recall of item 4. It was intended simply to compare these recall probabilities for the two conditions, with and without knowledge of structure, but due to an unforeseen difficulty a more complex analysis proved necessary.

III. Results

The raw data are summarized in Fig. 1. Although item 4 is recalled less often for the *K* (with knowledge) group, the difference—32 subjects against 36 who recalled this item under *I* (ignorance) conditions—is so small that its significance is lost. With these data the temptation is to accept the null hypothesis and assume that knowledge of structure has no effect. However, an odd point brought to light an important factor, which on reflexion seems obvious but which had nevertheless been overlooked. Turning our attention to the isolated item in the list in position 17 we see that the *K* group recalled this not less often but twice as often as the *I* group. In position 17 we should anticipate practically no difference, since in the original experiment the effect of isolating an item in this position was minimal (recalled by 7 subjects when isolated, and by 5 when massed).

In seeking for an explanation of this anomaly we are obliged to query the tacit assumption that the two groups were in fact comparable. Sampling errors are always with us, and in this case if we look at the total recall figures we see that while the *I* group recalled altogether 248 items, the *K* group recalled a total of 292. Testing this difference in a 2×2 table which includes the forgetting frequencies we find that it is significant at the $2\frac{1}{2}\%$ level on a two-tail test.



	26	17	11	36	16	2	9	6	13	6	17	16	6	7	2	8	7	8	6	29	Total
oup	28	12	14	32	18	6	16	10	13	9	14	15	12	11	9	11	14	9	11	28	248
roup	23.8	10.2	11.9	27.2	15.3	5.1	13.6	8.5	11.0	7.6	11.9	12.7	10.2	9.3	7.6	9.3	11.9	7.6	9.3	23.8	292
corrected																					248

Fig. 1. Effect of knowledge of structure. ○—○, ignorant of structure; ○- - - ○, with knowledge of structure; □, massed; △, corrected.

If there is a tendency for overall better recall by the *K* group then we should expect this to reduce the difference for item 4. The easiest way to allow for this factor of overall recall probability is simply to reduce all scores for the *K* group in the ratio 248:292. The result of this is to reduce the score for the *K* group in position 4 from 32 to 27.2. Incidentally, the recall score for the other group in the original experiment for whom the item in position 4 was massed, is raised slightly from 12 to 13.5 because their overall recall score was low; 221 in fact. Corrected values appear in Fig. 1 as triangles. If we compare the recall scores 36 and 27.2 for the *K* and *I* groups we find that the difference is significant at the 5% level on a two-tail test. Two objections need to be met before we can reject the null hypothesis with this level of confidence. First it may be argued that the groups differ in overall recall scores not because the samples were unfortunate, but *because* of the difference in treatment accorded them. This is not altogether fantastic, since knowledge of structure could decrease anxiety and lead to a more composed approach to the task. Even if this were true it would still be irrelevant for the purpose of testing the 'surprise' hypothesis. Suppose, to use a *reductio ad absurdum* argument, that

overall scores for the *K* group were increased to the point where, although item 4 no longer stood out in the serial position curve, it was nevertheless recalled in absolute terms more often than for the *I* group, we would still have proved the point that the better recall due to unexpected change is lost under *K* conditions. It is the relative and not the absolute recall score of item 4 that is of interest in assessing the surprise effect. Relative, that is, to the remaining items.

Furthermore, when we consider that the *I* sample came from the students of an evening institute, whereas the *K* sample was of university students and graduates, it is hardly surprising to find a higher recall ability in the second sample. Allowing then, that some sort of adjustment must be made for overall recall scores in order to make a comparison of the type we are interested in, are we justified in making a simple proportional correction throughout? We are, provided we are willing to make the assumption that all items are affected in the same proportions by any such overall effect. Although this seems a reasonable enough assumption to make, it may still be wrong. To guard against this danger another method of analysis is required.

Table 2. *Frequencies of recalling item 4 as a function of total number of items recalled*

<i>P</i>	<i>I</i> Group		Total items recalled	<i>K</i> Group		<i>P</i>	τ_b
	Including 4	Forgetting 4		Including 4	Forgetting 4		
1.00	1	0	1	0	0	—	—
0.50	1	1	2	0	2	0.00	+0.577
1.00	2	0	3	2	2	0.50	+0.500
0.77	10	3	4	2	2	0.50	+0.251
0.50	5	5	5	2	0	1.00	-0.378
1.00	6	0	6	7	5	0.58	+0.439
1.00	4	0	7	6	2	0.57	+0.316
1.00	3	0	8	7	1	0.88	+0.194
0.67	2	1	9	3	0	1.00	-0.448
1.00	2	0	10	2	0	1.00	0.000
—	0	0	11	1	0	1.00	—
Total subjects	36	10		32	14	$\Sigma\tau_b$	+1.451
	(a)	(b)		(c)	(d)	$\bar{\tau}_b$	+0.181
						$\Sigma\tau_b^2$	1.320

$$\tau_b = \frac{(ad - bc)}{\sqrt{(a+c)(b+d)(a+b)(c+d)}}$$

$$t = \frac{0.181\sqrt{8}}{\sqrt{0.165 - 0.033}} = 1.415 \quad \eta = 8 \quad P \approx 0.1 \text{ (one-tail)}$$

Although the groups are not equated in terms of overall recall scores, and it is impossible in principle to know beforehand whether or not they will be so equated, we can still overcome this difficulty by equating them along this dimension using the actual data. For example, we can take all the subjects of the *K* group who recalled a total of five items correctly and compare them with those of the *I* group who scored the same total. Our hypothesis is that a greater proportion of the *I* group than the *K* group will have recalled item 4.

An analysis of the data for all actual scores is given in Table 2. Comparisons for totals of 1 and 11 are impossible because no member of the *K* group recalled only one item and no member of the *I* group recalled as many as 11 items, so we lose two subjects, one from either group, by restricting our analysis to the common range.

A graph plotting the recall probabilities of item 4 against total items recalled for each total from 2 to 10 (Fig. 2) is too erratic to tell us much because of the small numbers

occurring in some of the cells. A moving average, grouping in fours, gives a smoother picture (Fig. 3). As predicted, the curve for the *I* group lies above that for the *K* group.

In order to make a valid statistical test the following procedure was adopted.* For each total in the common range the statistic τ_b was calculated. The formula is given in Table 2 and a full discussion is to be found in Kendall (1948). Without going into detail

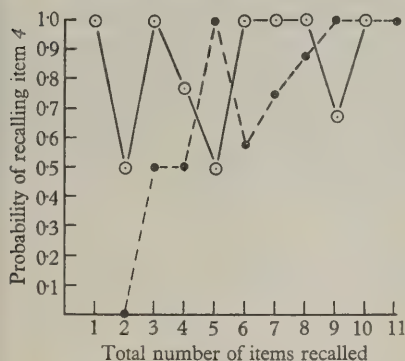


Fig. 2

Fig. 2. Recall probabilities of item 4 as a function of total number of items recalled. $\odot$ — $\odot$, *I* group; $\bullet$ — $\bullet$, *K* group.

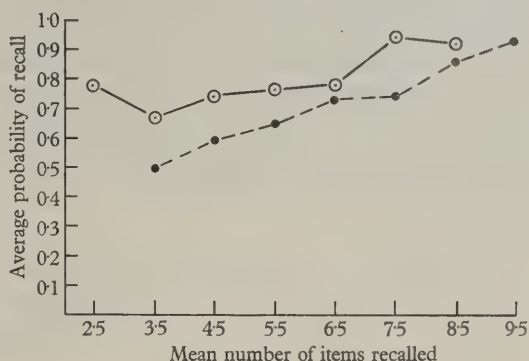


Fig. 3

Fig. 3. Moving average over four sub-groups. $\odot$ — $\odot$, *I* group; $\bullet$ — $\bullet$, *K* group.

we may note that τ_b will be positive when the graph for the *I* group in Fig. 2 lies above that for the *K* group, and negative when the opposite trend occurs. This measure is also partly weighted to allow for the number of subjects contributing to each sub-sample. So in essence we have a means of testing whether the curve for the *I* group lies above that for the *K* group. Testing against a theoretical mean of zero, a *t*-test gives $P \approx 0.1$ (one-tail), which is not very impressive. From a technical point of view a more satisfactory treatment is to use the Fisher-Yates exact test for 2×2 tables for each total compared, and combine the probabilities. This method gives $P \approx 0.07$.

IV. Discussion

Although the predicted reduction in the recall probability of the item in position 4 was found when the effect of overall recall scores was allowed for, the drop was by no means as marked as anticipated. If 'surprise' is the important factor and if previous instructions eliminate it then it is to be supposed that the recall of the item in position 4 under the *K* condition would be very similar to that where the same item is not isolated under the *I* condition. The actual difference between 32 and 12 is highly significant, and even if the corrected scores of 27.2 and 13.5 are used $\chi^2 = 8.27$ and $P \approx 0.004$.

Several possible explanations seem to offer themselves apart from sampling errors. The anomaly of the result was apparent some time before all the subjects had been run, and advantage was taken of this fact by asking the remaining subjects for retrospective reports. Two important points arose from these inquiries. In the first place the instructions were too complex to be grasped in detail and although most subjects realized that

* For the original idea behind this analysis I am indebted to Prof. L. J. Cronbach, and for the actual statistical technique to Dr A. R. Jonckheere and Dr R. Audley.

two types of material would appear the actual order was not mastered. Secondly, those subjects who grasped the fact that an item would be isolated at either end of the list reported waiting for these items to appear. In other words, the more effective the instructions were in eliminating 'surprise' the more they served to direct attention to the isolated items.

So, it would seem, *change* as such has attention-getting value whether expected or not, though the difference between the *K* and *I* groups suggests that unexpected change has additional attention-getting value. In the realm of sensory perception it is almost trite to refer to the effectiveness of change in securing attention, even to the extent of not noticing the tick of a clock until it stops. It is hardly surprising that in an experiment of this type the greater attention accorded an item introducing a change is reflected in a greater probability of its recall.

EXPERIMENT 2. THE RELATIVE EFFECTS OF ISOLATION AND STRUCTURAL CHANGE UPON RECALL

I. *Introduction*

It has been suggested by von Restorff (1933) that, in a temporal series, those items which are isolated, in the sense of being preceded and succeeded by items of a different type, will be better remembered because the figure-ground effect holds even when the material is presented item by item. It is supposed that this is in some way due to changes in the engrams, similar engrams becoming confused while the odd engram escapes.

More recent work by Green (1956) has shown that structural change is the more influential factor in affecting recall probabilities. In a list containing two critical items, only the first one is remembered significantly better when isolated than when surrounded by similar material. The suggested explanation was that while the first isolated item introduced a structural change, being the first item of a new type, the mixed nature of the list was well established by the time the second isolated item appeared. However, there was still a weak tendency for the second critical item to be remembered better when isolated than when massed, and this leaves us with the question as to whether this is a residual effect of structural change, or whether it indicates a small but genuine isolation factor. To settle this question it is necessary to devise an experiment which will separate these two factors.

II. *Experimental procedure*

In order to distinguish between the effects of isolation and structural change three groups were used, each being presented with a list that differed in structure, although the same 12 items were involved. These three lists are shown in Table 3. To control serial position and item-uniqueness effects the analysis is confined to the probability of recalling *CZ*, the item which occurs in the fourth serial position for all three groups. A perceptual paradigm of noughts and crosses is put alongside the lists to make the structural properties clearer.

In the change group item 4 is the first letter combination to appear and thus introduces a structural change. For the isolation group, item 4 not only introduces a structural change but is in addition isolated. So if isolation, as distinct from mere structural change, does facilitate recall then *CZ* should be remembered more often by the isolation group than the change group. This boils down in effect to asking whether or not the *succeeding* items have any influence upon the probability of recalling the critical item. If no difference is found between the two experimental groups, then in behaviourist serial interference terms, such as developed by Gibson (1940), the hypothesis becomes that while pro-active interference is considerable, retro-active interference is negligible.

The control group was added to provide a datum for the two experimental groups. In this list *CZ* is neither isolated nor is it the first letter item. It is predicted that both experimental groups will recall *CZ* better than the control group. If the structure of the succeeding items is irrelevant then there should be no difference between the two experimental groups.

The 167 subjects were students at a Technical College, nearly all males about 16 years of age. The material was presented to groups of 15–20 at a time by means of film strip and a projector in a semi-darkened room. Each item appeared for 3 sec. and was followed immediately by the next. Before the experiment proper started an informal chat put the subjects at their ease. They were assured that this was not an intelligence nor a personality test, and complete anonymity was guaranteed. Then formal instructions were given as follows: 'This is a memory experiment. I shall show you a series of items one at a time. After the series is complete you will have 1 minute in which to write on the slip provided as many of the items as you can recall. The order makes no difference. I am interested only in the total number of items that you recall correctly. Please attend carefully and avoid making any comments. Is there anything you are not quite sure about?' Nobody appeared to require more than the minute in which to record all they could remember.

Table 3. *Structure of lists*

Serial positions	Control group		Change group		Isolation group	
1	<i>17</i>	×	<i>17</i>	×	<i>17</i>	×
2	<i>RQ</i>	0	<i>53</i>	×	<i>53</i>	×
3	<i>94</i>	×	<i>94</i>	×	<i>94</i>	×
4	<i>CZ</i>	0	<i>CZ</i>	0	<i>CZ</i>	0
5	<i>68</i>	×	<i>RQ</i>	0	<i>68</i>	×
6	<i>TJ</i>	0	<i>TJ</i>	0	<i>21</i>	×
7	<i>79</i>	×	<i>BD</i>	0	<i>79</i>	×
8	<i>NL</i>	0	<i>NL</i>	0	<i>RQ</i>	0
9	<i>53</i>	×	<i>XP</i>	0	<i>TJ</i>	0
10	<i>BD</i>	0	<i>68</i>	×	<i>BD</i>	0
11	<i>21</i>	×	<i>21</i>	×	<i>NL</i>	0
12	<i>XP</i>	0	<i>79</i>	×	<i>XP</i>	0

III. Results

To give a general picture serial position curves for the three groups are shown in Fig. 4. A fair approximation to the usual bow shape is apparent except for position 4 in the case of the two experimental groups. While only 30% of the control group recalled *CZ*, some 60% of the isolation group and 65% of the change group recalled this critical item. The actual recall frequencies are given in Table 4 and apparently the difference between the control and experimental groups is highly significant. The difference between the two experimental groups is not marked and occurs in the opposite direction to that predicted by the isolation hypothesis. Examination of the overall recall scores, however, indicates that the groups were by no means matched in this respect. Random sampling errors seem to have placed the better memorizers in the change group and the worse in the control group. A similar difficulty was met in the preceding experiment and is dealt with in the same way. The result of this analysis is set out in Table 5.

This enables us to plot the probability of recalling item 4 as a function of overall recall scores for each of the three groups, as is shown in Fig. 5. Our hypothesis would lead us to predict that the graphs for the experimental groups will lie above that for the control group, since when overall recall ability is equated the probability of recalling item 4 should be greater under change and isolation conditions.

This general tendency is apparent, the graph for the control group only once going above either of the experimental groups. Using the same statistical treatments as were introduced in the first experiment we find that comparing the control and change groups

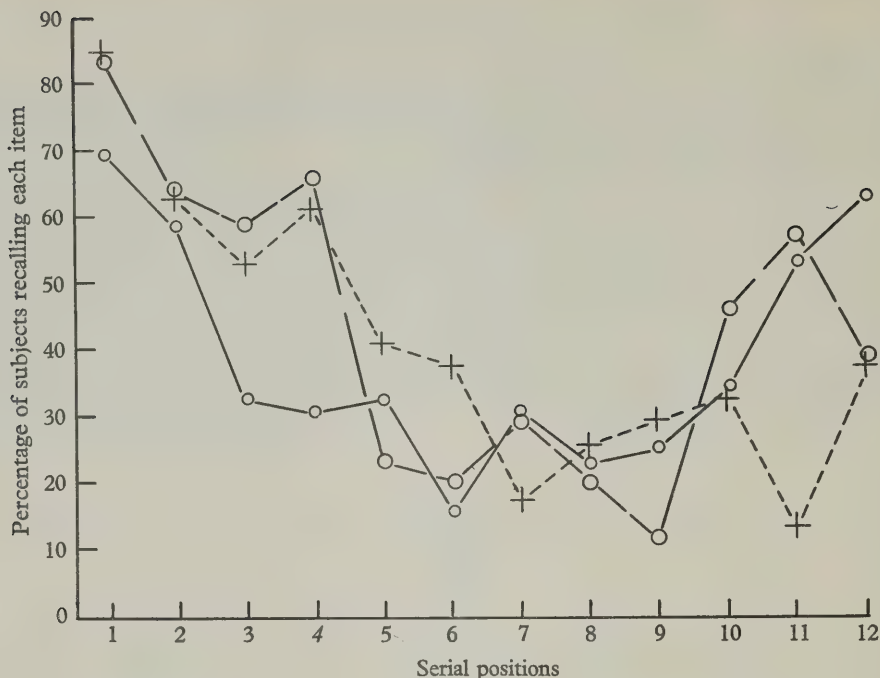


Fig. 4. Serial position curves. $\circ$ — $\circ$, control; $\circ$ — $\circ$, change; +---+, isolation.

Table 4. *Recall frequencies of item 4*

	Control group	Change group	Isolation group
No. of subjects	53	55	59
R4	16	36	36
F4	37	19	23
Recall (%)	30.2	65.5	61.0
Average no. items per subject	4.64	5.13	5.02

Control *v.* Isolation

$$\chi^2 = 8.62 \left(\frac{1}{27.4} + \frac{1}{31.6} + \frac{1}{24.6} + \frac{1}{28.4} \right)$$

$$= 10.6 \quad P \approx 0.001 \text{ (two-tail)}$$

R4, those subjects recalling item 4; F4, those subjects forgetting item 4.

Table 5. *Equating sub-groups for recall ability*

Total no. items recalled	Control group			Change group			Isolation group		
	R4	F4	P	R4	F4	P	R4	F4	P
1	0	2	0.00	0	0	—	0	1	0.00
2	0	3	0.00	1	3	0.25	0	2	0.00
3	1	8	0.11	3	5	0.38	3	4	0.43
4	6	8	0.43	7	4	0.64	5	5	0.50
5	0	9	0.00	7	2	0.78	7	6	0.54
6	3	3	0.50	6	3	0.67	14	4	0.78
7	3	4	0.43	6	1	0.86	5	1	0.86
8	1	0	1.00	3	1	0.75	1	0	1.00
9	2	0	1.00	2	0	1.00	0	0	—
10	0	0	—	1	0	1.00	0	0	—
Total	16	37		36	19		36	23	

R4, those subjects recalling item 4; F4, those subjects forgetting item 4.

by the first method: $t=2.32$; $\eta=7$; $P\approx 0.06$ (two-tail). For the control and isolation groups: $t=2.65$; $\eta=7$; $P\approx 0.05$ (two-tail). There is no discernible difference between the change and isolation groups beyond a faint trend of no significance in favour of the change group. The exact treatment leads to closely similar conclusions.

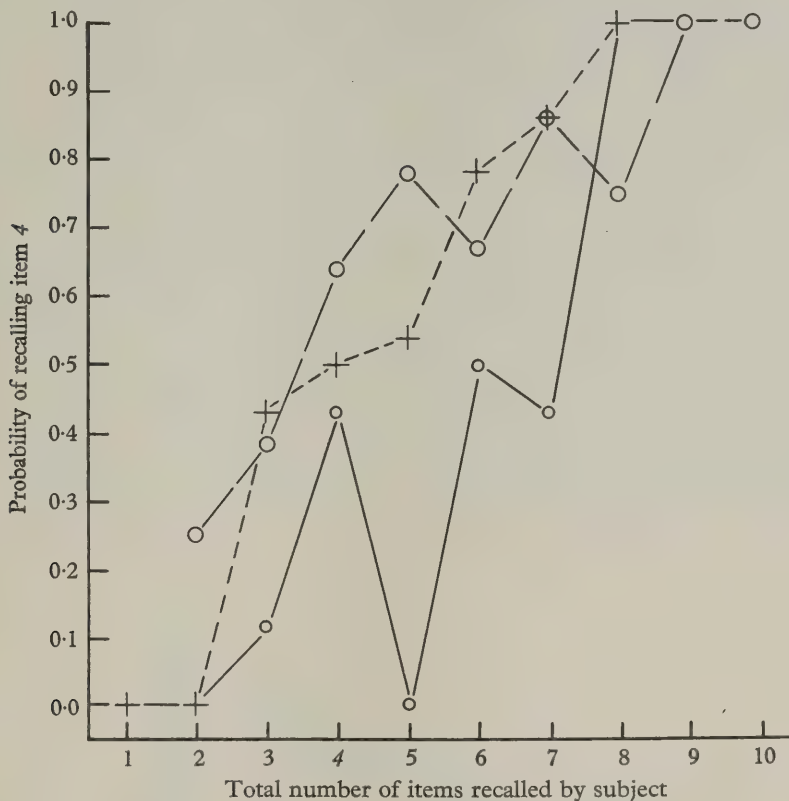


Fig. 5. Recall probability of item 4. o—o, control; ○—○, change; + - - +, isolation.

IV. Discussion

The obvious conclusion to be drawn from this result is that isolation in the von Restorff sense of the term has no effect. The relevant variable in this case appears to be structural change and only when an isolated item introduces such a change is it favoured in recall. So it would seem that earlier findings, including some of quite recent origin such as that of Kimble & Dufort (1955), which produced results favouring the isolation hypothesis, are an artefact resulting from the confounding of variables.

Another way of interpreting the data for CZ is in terms of pro-active interference. Structural change as defined in this experiment means introducing an item of a new type. So it could be argued that the main reason why the critical item is recalled less often by the control group is that CZ has been preceded by another letter item, RQ, which in some way either inhibits the perception of, or interferes with the engram of the critical item. For our present purposes, whatever the alternative explanation may be, the evidence is sufficient to discredit the isolation hypothesis.

REFERENCES

- GIBSON, E. J. (1940). A systematic application of concepts of generalisation to verbal learning. *Psychol. Rev.* **47**, 196-229.
- GREEN, R. T. (1956). Surprise as a factor in the von Restorff effect. *J. Exp. Psychol.* **52**, 340-4.
- KENDALL, M. G. (1948). *Rank Correlation Methods*, p. 35. London: Griffin.
- KIMBLE, G. A. & DUFORT, R. H. (1955). Meaningfulness and isolation as factors in verbal learning. *J. Exp. Psychol.* **50**, 361-8.
- VON RESTORFF, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychol. Forsch.* **18**, 299-342.

(Manuscript received 26 November 1956)

IMPOSSIBLE OBJECTS: A SPECIAL TYPE OF VISUAL ILLUSION

BY L. S. PENROSE AND R. PENROSE

(*University College, London, and Bedford College, London*)

Two-dimensional drawings can be made to convey the impression of three-dimensional objects. In certain circumstances this fact can be used to induce contradictory perceptual interpretations. Numerous ideas in this field have been exploited by Escher (1954). The present note deals with one special type of figure. Each individual part is acceptable as a representation of an object normally situated in three-dimensional space; and yet, owing to false connexions of the parts, acceptance of the whole figure on this basis leads to the illusory effect of an impossible structure. An elementary example is shown in Fig. 1. Here is a perspective drawing, each part of which is accepted as representing a three-dimensional rectangular structure. The lines in the drawing are, however, connected in such a manner as to produce an impossibility. As the eye pursues the lines of the figure, sudden changes in the interpretation of distance of the object from the observer are necessary. A more complicated structure, not drawn in perspective, is shown in Fig. 2. As this object is examined by following its surfaces, reappraisal has to be made very frequently.

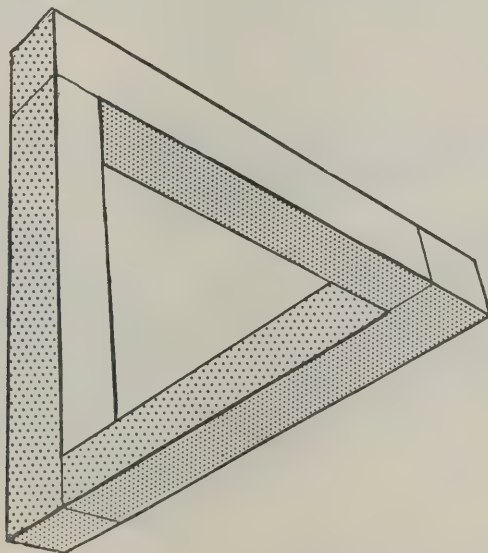


Fig. 1. Perspective drawing of impossible structure.

Another way of presenting the same type of illusion is to express the impossibility in terms of such a phenomenon as a continually descending or ascending path. The flight of steps drawn in Fig. 3 is an example of this. Each part of the structure is acceptable as representing a flight of steps but the connexions are such that the picture, as a whole, is inconsistent; the steps continually descend in a clockwise direction.

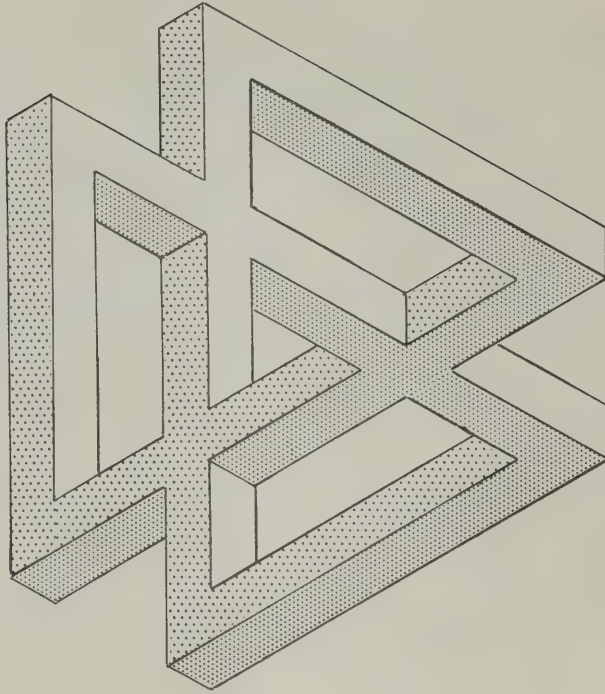


Fig. 2. Diagram of structure with multiple impossibilities.

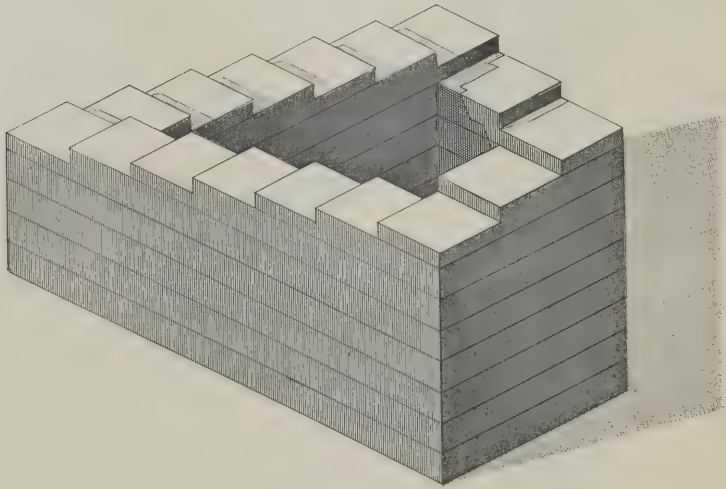


Fig. 3. Continuous flight of steps: shadowed drawing.

Actual objects suitably designed, when viewed from particular angles, can give exactly the same impressions as inconsistent drawings. A photograph of a model of this kind, apparently an impossible staircase, is shown as Fig. 4. Actually the extreme right hand step was much nearer to the camera and on a much higher level than the step which appears to be just above it. Illusions with a different purpose, constructed in a somewhat similar manner, have been discussed by Kilpatrick (1952).

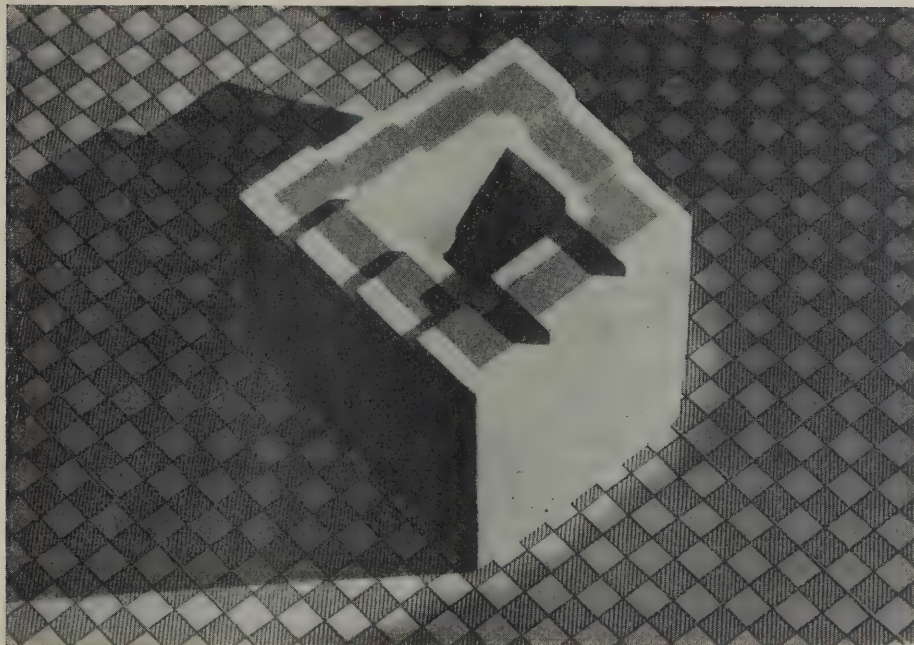


Fig. 4. Untouched photograph of continuous steps: the actual object is rectilinear and the surfaces of the steps are horizontal.

REFERENCES

- SCHER, M. C. (1954). *Catalogus 118*; Stedelijk Museum, Amsterdam.
- ILPATRICK, F. P. (1952). Elementary demonstrations of perceptual phenomena. In *Human Behavior from the Transactional Point of View*. Hanover, N.H.: Institute for Associated Research.

(*Manuscript received 30 November 1956*)

THE EFFECTS OF THE MENSTRUAL CYCLE ON TEST PERFORMANCE

BY MARY WICKHAM

It has been suggested that the menstrual cycle may represent a source of unreliability in the group testing of older girls and women. Data are presented for 4000 young women in which test scores are related to menstrual and health variables. It would appear that, during the Period Phase, there tends to be a slight depression of scores, other than in practical/mechanical performance, but this is too inconsistent and insignificant to warrant compensation in prediction, being no greater than the effects of the more common minor ailments.

When it is necessary to test women in large groups some of them will be passing through the physical disturbance of menstruation which may thus seem to present a problem additional to that of the usual minor ills. The aim of this investigation was to find out whether this disturbance influenced test performance, and if so to what extent. Two groups of women were used as subjects. Group A contained 3000 and group B 1000. Members of group A, all of whom had been in the A.T.S. for at least six months, were attending a Selection Centre with a view to changing their jobs. There they were interviewed, tested, and reclassified. The operation lasted from early February until late July so that weather conditions varied considerably. This group proved to be fairly representative of the service as a whole in terms of Progressive Matrices scores and educational background, but had more women of low medical category than was to be expected by chance (Table 1). Group B consisted of recruits who were first tested during a spell of extremely cold weather in February and March. They, too, were representative in terms

Table 1. *Distributions of both groups compared with a representative sample*

Selection Grade*	Group A (%)	Group B (%)	Representative sample (%)
(a) By general ability (on the Progressive Matrices)			
1 Top 10%	16.2	16.3	18.8
2 Next 20%	26.0	32.7	28.5
3 + Next 20%	23.6	24.3	23.0
3 - Next 20%	16.5	16.3	15.0
4 Next 20%	15.0	10.4	12.8
5 Bottom 10%	2.7	—	1.1
(b) By educational background			
Left school after 16, with School Certificate (or better qualification)	5.5	6.5	5.7
Left school at 16 without School Certificate	14.0	11.5	11.9
Left school at 15 or attended night school	12.9	15.5	18.0
Left school at 14 from top class	55.2	53.5	52.2
Left school at 14 but not from top class	12.4	13.0	12.2
(c) By medical category†			
A.W.	74.5	95.6	89.5
B.W.	11.4	4.4	9.5
C.W. and E.W.	14.1	—	1.0

* See Vernon & Parry (1949), p. 177.

† The medical categories used for the classification of the women were very similar to those used for the men. Thus A.W. was the classification for those of good physical standard; B.W. for those with some physical fault or handicap (frequently a slight defect in eyesight might account for this); while in the C.W. and E.W. categories the faults were more marked and serious.

of test scores and educational background, but had an unduly large proportion in high medical categories (Table 1).

Both groups of subjects did the following tests, all of which are fully discussed by Vernon & Parry (1949):

	Test	Items	Reliability
P. 1.	<i>Progressive Matrices</i> (Raven, 1939)	60	0.88
P. 2.	<i>Mechanical Comprehension</i> (Bennett)	50	0.85
P. 23.	<i>Arithmetic</i>	30	0.89
P. 4.	<i>Squares test</i> . This spatial test consisted of figures in which the testee had to draw a dividing line so that the two pieces formed would, if turned round, make a square	50	0.85
P. 14.	<i>Spelling</i> . Five incorrect and one correct version of each word were given	50	0.91
P. 21.	<i>Instruction test</i> . This test involved the comprehension of rather complicated instructions	150	0.92
P. 24.	<i>Mec. test</i> . This was a practical test involving the assembly of Meccano parts to a model with the aid of diagrams	7	0.77
P. 25.	<i>Verbal test</i> . This was based on synonyms and rhymes	104	0.97

Other information was collected under a number of heads, the procedure being different between the two groups. The data listed below were available.

Group A

When these women were at the Selection Centre their medical category was checked; they were given the whole battery of tests described above; they were interviewed by a selection officer, a medical officer, and if, in the opinion of either of these officers, it was thought desirable, they were also interviewed by a psychiatrist. In the course of this procedure the following information was obtained for each woman: (i) scores on the whole test battery described above; (ii) re-checked medical category; (iii) educational background; (iv) service and previous civilian employment; (v) date of the onset of the immediately previous menstrual cycle; (vi) average length of the menstrual cycle; (vii) average duration of the menstrual period; (viii) incidence and severity of dysmenorrhoea; (ix) a rating on mental stability (this rating was given by the selection and medical officers and in cases where they suspected serious difficulties the opinion of the psychiatrist was included); (x) the original Matrices grade which had been obtained when the woman joined the service, taken at least six months earlier and at random in so far as the menstrual phase was concerned.

Group B

These women were tested with the whole test battery and then retested with the same battery in six weeks time. Each was interviewed and medically examined and herself filled in a questionnaire. During the course of this procedure the following information was obtained for each of these women: (i) scores on the whole test battery; (ii) retest scores for the same; (iii) medical category; (iv) educational background; (v) date of the immediately previous and of the next expected menstrual period (this was obtained at both the time of the test and retest); (vi) any feeling of dysmenorrhoea at either the test or retest; (vii) any other factor which they felt to be having a disturbing influence at either test or retest (the common cold and chilblains were the handicaps most frequently named, being a reflexion of the severity of the weather at the time of the investigation).

In previous work on this subject the investigations, such as those of Sowton & Myers (1928), had been concerned with the performance of individuals over a series of menstrual

cycles. A tendency to depression in functional activity shortly before and during menstruation and a heightening in the middle of the inter-menstrual phase, had been noted by such writers as Bedale (1928). The findings have been far from consistent. A general tendency for menstruation to depress test scores had been claimed. Some divergence between the effect in the performance of the same individual in different tests had also been noted by Sowton & Myers (1928), the scores in the tests of an intellectual type being depressed while those of manual skills were in some cases even increased. In most of this previous work the limitations and difficulties inherent in the tracing of the performance of individuals over many cycles were stressed. In this investigation some of these difficulties were avoided. A wide range of tests was used, involving about two hours testing time; such could obviously not be done in daily repetition. The tests were given on two occasions only, but to large numbers. The women tested did not think the information they gave about their menstrual cycle of much import, as it seemed merely incidental to the main aims at the testing.

The results in the second series of testing were compared with those of the first to see to what extent they varied beyond expectation, when the subjects were passing through some disturbed phase in the menstrual cycle.

The data obtained from group A were dealt with as follows. Length and duration of the menstrual cycle were considered. No less than 80% of this group gave the usual length of their menstrual cycle as 28 days, 8% gave a shorter and 12% a longer cycle than this, varying from 14 days on the one hand to 43 on the other. There was a marked tendency for the cycle to be considered in multiples of 7 and given in weeks rather than days, thus there were peaks rising to 30 individuals at 14 days, 120 at 21 days, 70 at 35 days and a minor peak at 42 days.

It is obviously difficult to assess the actual length of 'period' with accuracy, as even within the experience of one individual there are fluctuations. Some 94% of this group quoted their 'period' as lasting from 3 to 7 days, over 800 women giving 5 days as the usual duration. Of the remainder, 6%, some claimed a 'period' as short as 1 day, while there was one case in which it lasted 14 days.

Incidence of dysmenorrhoea

During the course of the interviews it was found that 24.4% of the group experienced some pain immediately before or during the 'period'. For the majority the pain was felt during the first or the second day. These figures are not very different from those of other writers; thus Sturgis (1923) found that 30.6% of over 2000 women employed in a store suffered some slight handicap while 4.4% found it serious. The incidence of dysmenorrhoea was greatest in the lower medical categories and in the more strenuous jobs, thus it was significantly more marked in the operational members in anti-aircraft than those occupied in the more sedentary jobs of clerking and signals. This is shown in Table 2.

Table 2. *Significance of differences in the percentage incidence of dysmenorrhoea*

- (a) By medical category: A.W. 23.28%, B.W. 26.20%, C. and E.W. 29.64%
% difference: A.W. - C. and E.W. 6.36; standard error 2.78
- (b) By job: operational nos. 27.0%, cooks 25.1%, clerks and signals 21.9%
% difference: operational nos. - clerks and signals 5.1, standard error 2.5
- (c) By normality of cycle and period: normal 19.54%, abnormal 29.26%
% difference: normal - abnormal 9.72, standard error 1.99

Owing to the great variation in length of cycle (noted above) a subdivision was made into groups giving the *Normal Cycle* (28 days), a *Short Cycle* (20, 21 or 22 days) and a *Long Cycle* (34, 35 or 36 days).

Two outstanding phases within the menstrual cycle have been shown by such writers as Hoskins (1941) and Benedek & Rubenstein (1939). In mid-cycle, the build-up of oestrogen and progestin culminates in ovulation. Towards the end of the cycle, the destructive forces are in control and they culminate in the menstrual period. Since the stimulating build-up and the following break-up take place gradually there are two other intervening phases of 'rest'. From 4 days before to 4 days after the onset of menstruation would seem to cover the period when the destructive and depressing phase is of greatest influence, and from the 11th to the 18th day would cover the ovulation, when the stimulation would be at a maximum in those women who had the normal length of cycle, and the normal length of period. The mid-cycle 'ovulation' phase was of equal interest to that of menstruation as it might have had a stimulating effect on the test scores. The nomenclature used by Benedek & Rubenstein (1939) was adopted, namely the 'ovulation' and the 'menstrual' phase and the two intervening phases.

Bearing these complexities in mind, and in view of the relatively large numbers of women whose results were available, it was decided to minimize the chance variables by including in this subgroup only those quoting a normal length of cycle and period. There were 1335 for whom this was so and for whom all other relevant information was available. In a few cases some of the earlier test results were not known and they had to be excluded. Most of the relevant data were available for 1525 cases.)

The 'short' and the 'long' cycle subgroups were similarly considered, the period phase being taken as from 3 days before until 3 days after the onset of menstruation, and from 5 days before to 5 days after the onset for the two subgroups respectively. There were seventy-nine cases claiming a 'short' cycle and fifty-five who claimed a 'long' cycle. As the numbers here were smaller than in the normal group, and there was little evidence for placing the 'ovulation' phase, the results when grouped for 'phases' were considered in relation only to 'period' and 'non-period' phases.

The cases whose results did not fall into one of these subgroups were ignored for the next part of the analysis. It was felt that more would be lost than gained by including all the intervening and outlying cases as the variables involved (such as abnormal length of cycle and duration of period) might well mask rather than elucidate any influence of the menstrual cycle upon the test results.

The 'normal' cycle subgroup of 1525 cases was then sorted into 'day' groups so that all the women who were at the second testing on the same day within the menstrual cycle were put into the appropriate 'day' group. This was done for each of the 28 days of the cycle. The mean of the 1525 cases was found for each of the tests given. The average deviation of each of the 'day' subgroups was found for each of the tests. Thus the *actual deviation* for each day within the menstrual cycle was obtained for the eight tests described on page 35.

The results of the Matrices test, which each of these women had taken at least six months earlier, had been scored at random, in so far as menstruation was concerned. The average score of each of the 'day' subgroups was again found, and then the deviation of each 'day' subgroup from the average of the 1525 (forming the whole group) having a normal cycle. Thus the expected deviations were obtained. For tests other than the

Matrices, the method given by G. W. Snedecor (1940) was followed. The inter-correlation of each of the tests with Matrices was obtained and the reliability of all the tests was known (see page 35). Thus the expected deviations were obtained.

The results of those falling into the phases as regards the second testing (i.e. the Period, the Ovulation, and the two intervening phases of rest as described on page 37) were combined to give the observed deviations for the 'phase' subgroups. These were then considered against the expected deviations.

If the different phases affect performances on these tests differently, we can expect that the average performances of different phase groups will vary. On the other hand, if we find that average performances vary, we cannot infer that this is due to differences in phase, since the normal standard of performance of the individuals in each phase group may be different. There are, thus, two possible explanations of any observed differences between the phase groups. Knowing the grades on the original testing with Progressive Matrices, which are independent of phase, we can estimate how far the phase groups differ in normal standard of performance. We can thus establish expected differences between

Table 3. *Significance of the excess of differences between the observed and expected performances*

Test	<i>F</i> value	<i>P</i>
S.P. 2. Mechanical Comprehension	1.7	>0.05
S.P. 23. Arithmetic	2.08	>0.05
S.P. 4. Squares	1.33	>0.05
S.P. 14. Spelling	1.65	>0.05
S.P. 21. Instructions	3.80	<0.05 >0.01
S.P. 24. 'Mec' test. Practical models	2.39	>0.05
S.P. 25. Verbal	1.32	>0.05
S.P. tests (sum of scores for all tests)	1.04	>0.05

the phase groups and measure how far the observed differences exceed the expected differences. The excess can be attributed to the effects on performance of different phases in the cycle. The error variance of the measurements can be derived from similar measurements comparing observations with expectations for separate day groups within the same phase group, and an *F* test can be applied to measure the significance of any excess found.

This preliminary review was made to see which of the tests warranted further analysis. The *F* values obtained for the significance of the variance 'within' as compared with the variance 'between' the four phases are given in Table 3.

S.P. test 1, the Matrices, is omitted from Table 3 as it was decided to make a more detailed analysis of this test. From the results it was concluded that the Instructions test also warranted further analysis. The test results of the women who were in the 'short' and 'long' cycle subgroups were analysed in the same way. Although slightly lower scores than expected were obtained on most of the tests by the women in the 'Period' phase, in no case was significance reached, it being most nearly approached in the Instructions test.

At this point it seems to be worth noting that the incidence of dysmenorrhoea was greatest in those who had abnormal cycles (see Table 2). For the group as a whole (all 3000), those claiming a normal cycle and period were compared with those falling outside these limits in either direction for either cycle or period in relation to the incidence of dysmenorrhoea. Since a very significant difference is found, the test results of the abnormal groups might be expected to show the greater disturbance in score.

There is a tendency for a depression of score on the paper and pencil tests to occur for those in the 'period' phase. This tendency is reversed in all groups, for the practical lec. test, and in the Arithmetic and Instructions tests for the group with an abnormally long cycle.

After this preliminary work, a more exhaustive analysis was undertaken for the Matrices and the Instructions tests. The Matrices was chosen because the initial test grades were available. The test-retest correlation was 0.790. The Instructions was chosen because, in the preliminary work, the difference in score had most nearly reached significance (see Table 3).

Table 4. *Analysis of variance of the 'days' within the menstrual cycle (Matrices test)*

Source of variance	Sum of squares	Degrees of freedom	Mean square
Total variance on retesting	1194.244	1334	—
Accounted for by regression of retest on test	744.350	1	—
Accounted for by differences in performance on different days	13.396	27	0.496
Residual variance	436.498	1306	0.334

$$F = 1.48, P > 0.05.$$

Table 5. *Analysis of variance due to menstrual phase (Matrices test)*

Source of variance	Sum of squares	Degrees of freedom	Mean square
Accounted for by differences between			
(a) Phases	0.7878	3	0.263*
(b) Days in the same phase	12.609	24	0.525†

$$* F = 1.27, P > 0.05.$$

$$† F = 1.57, P < 0.05 > 0.01.$$

Table 6

Source of variance	Sum of squares	Degrees of freedom	Mean square
Total variance on retesting	5247.1174	1524	—
Accounted for by regression			
(i) Of Instructions on Matrices	2085.8822	1	—
(ii) Differences within phases	10.0354	3	3.35
(iii) Differences between days within the same phase	35.5885	24	1.48
Residual variance	3115.6113	1496	2.08

$$F \text{ (ii) } 1.61, P > 0.05.$$

$$F \text{ (iii) } 1.40, P > 0.05.$$

The differences between the observed and the expected scores of the women whose retest fell on each of the 28 days of the menstrual cycle were first obtained separately and analysed (see Table 4). In spite of the large numbers involved the differences are only on the borderline of significance. The analysis was then carried a step further. All the women tested during the same 'phase' were grouped together (see Table 5). This table shows that such differences as occur between observed and expected mean retest scores are irregular and not *significantly* associated with a *particular day or phase* within the menstrual cycle.

A similar analysis was then made for the Instructions test. The mean scores were estimated from the results of the original testing on the Matrices. The correlation between the Matrices and Instructions tests for this sample was 0.631 (see Table 6).

The significance indicated for the Instructions test in the preliminary analysis is not confirmed.

It can then be concluded from the results of this group of women that no advantage would be gained by taking into account the day within the menstrual cycle on which the test was given for predictive purposes.

In group B the data were dealt with as follows. Since the test and retest results for each test of the battery were available for each of the 1000 women in this group, a simple method was available. There was a 6-week interval between the first and second testing, and so the results obtained for each woman were for some different day within the menstrual cycle. If the menstrual phase influenced scores adversely, then the women in this phase at the first testing would not, at that time, do themselves justice. At the retest these women, not being in this phase of expected depression, would show their real capacity. Therefore the differences between their test and retest scores would be *greater* than the differences between the test scores of the women who were not in the menstrual (period) phase at either testing or retesting. Conversely, the women in the menstrual (period) phase at the time of retest would be handicapped at the second testing, and the difference between their test and retest scores would be less than those of the women who were not in this phase at either time.

Table 7. *The means of the differences in test scores*

Test	(b - a)	S.E.	(c - a)	S.E.
	expected positive		expected negative	
S.P. 2. Mechanical, paper and pencil	+0.109	0.642	1.440	0.814
S.P. 23. Arithmetic	0.943	0.773	-0.559	1.106
S.P. 4. Squares	-0.970	0.665	1.176	0.730
S.P. 14. Spelling	-0.743	0.574	0.075	1.025
S.P. 21. Instructions	-0.139	1.228	-0.377	2.070
S.P. 24. Mechanical, practical	-3.55	1.50	-2.325	2.00
S.P. 25. Verbal	0.760	0.733	0.191	0.909

In order to examine these two hypotheses the means of the differences between the test and retest scores were obtained for: (a) those outside the menstrual phase at both test and retest; (b) those in the menstrual phase at the initial testing; and (c) those in the menstrual phase at retesting. The differences of the means for the above groups are given in Table 7. In view of the above hypotheses the value for (b) is expected to be greater than that for (a), and that for (c) is expected to be less than that for (a).

As can be seen, all differences are insignificant, except in one case (S.P. 24), where the difference is in the direction contrary to that expected.

As was noted on page 35, the common cold and chilblains were the most frequently quoted cause of disturbance in this group other than dysmenorrhoea. At the time of retesting no fewer than 177 reported severe colds.

In order that the effect of these minor ills could be compared with that of dysmenorrhoea, parallel calculations were made for the results of the Instructions and the practical Mechanical tests, as in the above work they had shown the most nearly significant results (see Table 8). Again the values found give no consistently significant trend. It is of interest to note that here there is no tendency for the Mechanical test to behave differently from the Instructions test as there was in the comparison where dysmenorrhoea was the handicap under consideration.

Table 8

	Instructions	Mechanical
Showing the differences between the means of the scores of those:—		
Handicapped by cold at the test—	- 0.765	0.550 (positive value expected)
The not-handicapped		
S.E.	2.05	2.09
Handicapped in the retest—	- 0.882	- 1.303 (negative value expected)
The not-handicapped		
S.E.	0.818	0.952
Handicapped by chilblains at retest	- 1.794	- 1.269 (negative value expected)
The not-so-handicapped at the retest		
S.E.	1.462	1.551

CONCLUSIONS

(1) In the results of the 3000 women in group A there was a tendency for depression of score in the Period Phase. This was noted for just over 70 % of the performances of the subgroups for the eight tests considered. (2) Half the exceptions to this tendency occurred in the results for the practical Mechanical test where the reverse was found, the women doing better during the Period Phase. (3) The depression of scores did not however reach a level of consistent significance. (4) The results of the 1000 women in group B were marked by inconsistency and a lack of significance. The influence of minor common ills such as colds and chilblains seemed equally likely to cause a fall in score. (5) The scores in the practical Mechanical test tended to be higher during the Period Phase, as for the first group. The scores in this test when the testees were under the handicap of colds, etc., were lower, behaving in the same way as the scores in the other tests.

Finally, it was concluded that no practical advantage would be gained by taking into account the day within the menstrual cycle on which a test was given for predictive purposes.

It will be appreciated that in an investigation of this kind the writer is grateful to the many officers and other ranks who gave and collected the data from which the conclusions were drawn, and above all to Dr Edith Mercer who made the investigation possible, and to Mr Patrick Slater and Prof. P. E. Vernon for their help and advice upon the statistical methods used.

REFERENCES

- BEDALE, E. M. (1928). Experimental study of the menstrual cycle. Its relation to general functional activity. *Rep. Med. Res. Coun. Industr. Res. Bd.*, no. 45.
- BENEDEK, T. & RUBENSTEIN, B. B. The correlations between ovarian activity and psychodynamic processes. *Psychosom. Med.*, 1, 244, 461.
- Bodily disturbances during the menstrual period. In *Practitioners' Encyclopedia of Midwifery and the Diseases of Women*. London: Frowde and Hodder and Stoughton. 1921.
- HOSKINS, R. G. (1941). *The Glands and their Functions*. London: Routledge & Kegan Paul.
- SNEDECOR, G. W. (1940). *Statistical Methods*, chapter 12. Iowa State College Press.
- SOWTON, S. & MYERS, C. S. (1928). Experimental study of the menstrual cycle. Its influence on mental and muscular efficiency. *Rep. Med. Res. Coun. Industr. Res. Bd.*, no. 45.
- STURGIS, M. C. (1923). Observations on dysmenorrhoea occurring in women employed in a large departmental store. *J. Industr. Hyg.* 5, 53.
- VERNON, P. E. & PARRY, J. B. (1949). *Personnel Selection in the British Forces*, p. 177. London University Press.

(Manuscript received 4 December 1956)

A 'CRESPI EFFECT' IN MALE IMBECILES

BY N. O'CONNOR AND G. S. CLARIDGE

*Medical Research Council, Social Psychiatry Research Unit,
Institute of Psychiatry, London*

In order to test the effects on performance of an interchange in incentive level forty male imbeciles were divided into four equal matched groups on the basis of two initial trials on a simple dexterity task. For a further ten trials two groups worked under non-incentive conditions and two under incentive conditions, in which they were encouraged to reach a goal based on their previous performance. On the last trial one incentive group was changed to control conditions and one control group to incentive conditions, the other two groups continuing as before. The experiment was continued for a further seven trials.

Results indicated that groups given the incentive from the beginning of learning separated significantly from those receiving no incentive. When one of the latter groups was given an incentive, however, its performance increase showed a typical Crespi type 'elation' effect, performance reaching, and maintaining, a level greater than that of subjects who had received the incentive from the beginning of learning. The effect was found to be a group trend rather than a response typical of all subjects.

A comparable decrease in incentive level was found to have little marked effect on performance level, which established itself at a point intermediate between the performance levels of subjects who continued to receive an incentive and of subjects who had received no incentive at all. This finding confirmed earlier work with defectives, showing that incentives would have relatively permanent effects on the attitude or set, and hence on the performance level of imbeciles.

I. INTRODUCTION

It is generally agreed by learning theorists in both the animal and the human fields that changes in level of motivation will normally result in corresponding changes in performance level. The precise nature and extent of these performance changes have, however, been the subject of some controversy, particularly among comparative psychologists.

In the final treatment of his theory Hull (1951) gave specific consideration to this problem and derived a theorem to cover the effects of changes in the incentive component (K). He says:

Other things constant, an abrupt shift in the incentive used during a maze-learning process will be followed first by a major shift in reaction potential and then by two or more progressively smaller shifts on successive trials, the series constituting a rapid learning process of the exponential variety, culminating in the course that the sE_R would have followed had the new incentive been operating continuously from the beginning of learning.

In deriving this theorem Hull (1951) makes use, somewhat paradoxically, of the work of Crespi (1942), who reports a very meticulous and elaborate study in which interchanges in a number of food amounts were made, the drive variable being carefully controlled at all stages of the experiment. Using speed of running as a measure of response, Crespi demonstrates changes in performance level which are greater than can be predicted from Hull's theorem. Thus, it was found that when a larger amount of food was introduced after previous training under a smaller amount, performance level increased to a point *above* that of animals who had received the larger incentive amount from the beginning of training. Crespi calls this excess increase an 'elation' effect and also records a corresponding 'depression' effect in animals shifted from larger to smaller amounts of food.

Similar contrast effects are also reported by Zeaman (1949) using latency as a measure of response.

While Hull explained what he himself termed the 'Crespi effect' as being merely a temporary, emotional phenomenon, Crespi (1944) attaches more significance to it. He suggests that the 'elation' effect, for example, is due to an increase in emotional drive as a result of the discrepancy between expectation, built up as a result of experience of a smaller incentive, and the sudden achievement of a larger incentive.

Much of the work carried out on changes in motivation in human subjects, which has tended to proceed independently of similar work in the comparative field, has been carried out with knowledge of results as the incentive situation. Mace (1935) and Book & Norvelle (1922) found rises and drops in performance when subjects were changed from no knowledge to knowledge of results and vice versa, but the results of Mackworth (1950) are not entirely consistent with these findings. He found that the introduction of knowledge of results produced a significant fall in the error score of radar watchers, but that the reverse procedure had no significant effect on performance.

In studies with humans differences in the kind of incentive and experimental task employed by investigators, as well as the failure to provide adequate control groups, have led to some uncertainty about the precise effects on performance of changes in incentive level, although on the whole, as pointed out by Ammons (1954), results have tended to corroborate the predictions of Hullian theory, the 'Crespi effect' not having been demonstrated in human subjects.

II. THE EXPERIMENT

(i) *Aims of experiment.* Previous work with imbeciles (Gordon, 1953; Claridge, 1956) had suggested that under certain conditions of motivation the 'elation' effect reported by Crespi (1942) might occur. It might be expected, for instance, that under conditions where subjects had been made accustomed over a relatively long period to performance on a monotonous routine task without incentive, a change in these conditions would produce the kind of 'emotional boost' proposed by Crespi (1944).

The outcome of a change in the reverse direction—that is, from more to less favourable incentive conditions—was, as in the case of normal human subjects, less easily determined. It was possible, in fact, that little change in performance would occur, for results reported by O'Connor & Claridge (1955) suggested that favourable incentive conditions early in practice would have effects on the performance of imbeciles which were persistent for periods of at least a year.

In order to throw some light upon these problems an experiment was designed which would provide for systematic increases and decreases in the level of motivation of imbeciles, with adequate control groups for these changes. It was hoped, thereby, to make a more precise comparison than has hitherto been possible with the current approach of comparative learning theorists to this problem.

(ii) *Subjects.* Subjects were selected from the imbecile population of the small branch institution of a large hospital for mental defectives. Forty patients between the ages of 18 and 45 were finally chosen, excluding those who were blind, epileptic, or severely crippled. Estimates of I.Q., made by means of the 'Draw a Man' test and scored according to the instructions of Goodenough (1926), confirmed the diagnosis of imbecile in all cases. Mean I.Q. for the 40 subjects was 41.02, S.D. 7.55 and mean age was 29.20, S.D. 7.71.

(iii) *The task.* The task used was a routine motor task termed the Nail Frame and devised by Gordon (1953). It consisted of a frame ($8 \times 10 \times \frac{3}{4}$) covered with a sheet of perforated zinc containing 82 holes to the square inch. A column down the shorter side of the board consisted of exactly 60 holes and subjects were asked to insert small escutcheon pins ($\frac{7}{16}$ in. in length) into the holes, working vertically down a column and completing it before starting the next. Each subject was given a small tray, containing a large supply of pins, which was placed at the side of the board appropriate to the handedness of the subject. Subjects were required to work for 60 minutes each day, this constituting a trial. The number of columns and the number of pins inserted in a partially completed column were easily calculated at the end of a trial by means of a graduated scale along the edges of the frame.

(iv) *The incentive.* The incentive condition used for the experiment was termed Goal with Encouragement, which had been shown (O'Connor & Claridge, 1955) to be maximally effective with male imbeciles. Under this incentive subjects were encouraged at intervals to reach a goal based on their previous performance. Achievement of the goal was praised and failure sympathetically observed by the experimenter.

(v) *Design and procedure.* Four groups of subjects were required and, since they were to be matched, two initial trials were given to form a basis for matching. On two successive days the 40 subjects were taken in groups of 10 and given an hour's trial on the Nail Frame, subjects simply being told to work as hard as they could. Two scores were, therefore, obtained for each subject, and on the basis of the mean of these scores subjects were assigned to one of four groups. The 40 subjects were ranked in order of initial performance level and split off into quadruplets, a member of each quadruplet being allotted at random to one of the four groups, which were designated as follows:

Goal-Goal (GG)	10 subjects
Goal-Control (GC)	10 subjects
Control-Control (CC)	10 subjects
Control-Goal (CG)	10 subjects

For the next 10 days (trials) of the experiment two groups (Goal-Goal and Goal-Control) worked under incentive conditions, which were as follows. *Goal with Encouragement conditions.* At the beginning of each trial a red nail was inserted in the board of each subject at a point which represented a 1% increase on the highest of his three previous scores. At the beginning of each trial subjects were told; 'Let's see you all get your goals today. You worked very well yesterday, but see if you can work even harder today.' Every quarter of an hour, they were told: 'Come on, boys, keep working hard. You are doing very well at the moment. Let's see you get your goals.' In addition to the group encouragement individual encouragement was given to each subject during the course of the trial. At the end of the trial those who succeeded in reaching their goals were told: 'That was very good; see if you can get your goal again tomorrow.' Those who failed were told: 'Bad luck; you did very well, anyway, and you'll get your goal tomorrow.'

For the first part of the experiment the other two groups (Control-Control and Control-Goal) worked under the following conditions. *Control or non-incentive conditions.* At the beginning of each trial subjects were told, 'work as hard as you can'; and at the end, 'you worked very well today; I'll see you all again tomorrow at the same time'.

At the end of this period, i.e. on the 13th trial of the experiment, a change in incentive conditions was made. One incentive group (Goal-Control) was changed to non-incentive conditions, the other incentive group continuing as before. At the same time one non-incentive group (Control-Goal) was changed to conditions of Goal with Encouragement, the other non-incentive group remaining on Control conditions. The experiment was then run for a further seven trials, at which point it was terminated.

III. RESULTS

In order to convert the data into a form which would give measures of relative improvement and performance level, a linear regression technique was applied to the results, treating the pre-change period (trials 3-12) and the post-change period (trials 13-20) separately. The regression coefficients for linear regression ($\text{score} = a(\text{trials}) + b$) for each individual are shown in Table 1, while Fig. 1 shows the group mean scores for each trial plotted as a graph.

The treatment of results in the periods prior to and following the change in incentive level will be considered separately in turn.

The pre-change period. It is apparent from Fig. 1 that at the point where the change in incentive was made the two incentive groups were clearly differentiated from the two non-incentive groups. This was found to be due, not so much to an increase in the over-all rate of improvement during this period, as to an abrupt increase in performance level when the incentive was introduced at the beginning of the experiment. Once the Goal-Goal and Goal-Control groups were established at this level the rates of improvement for

all four groups were closely similar, as shown by the following group means for the 'a' coefficients during this period: Goal-Goal, 8.09; Goal-Control, 7.95; Control-Control, 7.27; Control-Goal, 7.60.

An Analysis of Variance confirmed that there was, in fact, no significant difference between the four groups in rates of improvement, F being 0.112. The superiority in level gained by Goal subjects on the first trial under these conditions was thus maintained throughout the pre-change period as revealed in the marked separation of the two pairs of groups. An Analysis of Variance computed in individual performance levels for the 12th trial indicated that there was a significant difference between incentive conditions at the end of the pre-change period ($F=5.965$, significant at 0.05 level). Although the significance level was a relatively lenient one, the clear over-all group trends and the way

Table 1. *Regression coefficients ('a' and 'b') for each subject in each group.*

Group	Subject	Trials 3-12		Trials 13-20	
		'a'	'b'	'a'	'b'
Control-Goal	1	9.9	738	-15.7	994
	2	20.0	612	-8.0	837
	3	0.8	744	4.7	981
	4	14.8	516	0.3	724
	5	3.7	596	1.0	767
	6	-19.6	674	7.1	668
	7	-2.2	346	-1.3	512
	8	27.3	417	5.5	662
	9	7.1	351	12.5	520
	10	14.2	326	5.9	528
	Mean	7.60	532.0	1.20	719.3
Control-Control	11	-3.6	862	4.7	851
	12	9.1	661	-5.0	833
	13	14.2	669	2.3	810
	14	6.6	580	-11.9	684
	15	-0.4	514	-10.4	482
	16	7.8	466	6.3	496
	17	16.5	424	-1.6	561
	18	28.1	357	-5.8	577
	19	-0.4	284	7.0	234
	20	-5.2	453	2.3	424
	Mean	7.27	527.0	-1.21	595.2
Goal-Goal	21	28.5	923	2.8	1138
	22	5.6	639	4.8	639
	23	4.4	672	-1.3	747
	24	11.8	623	1.0	824
	25	3.5	712	-9.6	845
	26	4.3	421	-11.2	500
	27	-4.1	430	-10.3	401
	28	24.2	480	-8.7	740
	29	-6.0	545	32.1	426
	30	8.7	420	10.9	486
	Mean	8.09	586.5	1.05	674.6
Goal-Control	31	11.1	778	-4.5	896
	32	-15.3	939	-6.3	815
	33	15.0	724	-4.9	865
	34	4.2	564	5.6	574
	35	5.9	619	11.6	644
	36	9.2	622	-4.8	688
	37	16.4	430	-2.3	574
	38	2.5	445	-2.9	483
	39	10.3	499	27.7	499
	40	20.2	214	-0.5	372
	Mean	7.95	583.4	1.87	641.0

in which the two Goal groups had kept together in separating from the two Control groups suggested that the differences were not due to chance but to a real effect of Goal with Encouragement.

The situation with regard to performance level, of the two pairs of groups at the point immediately prior to the interchange in incentive conditions indicated, therefore, that a critical test of the effect of this change could be made.

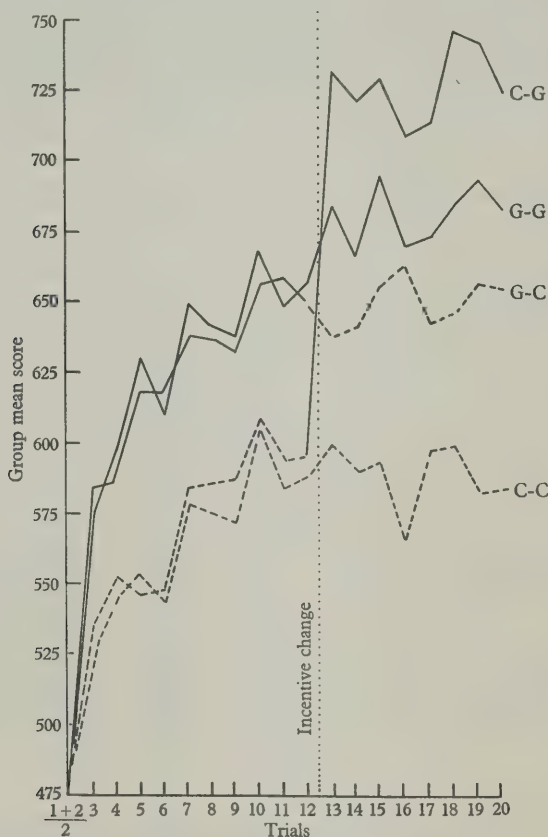


Fig. 1. Group mean scores for each trial. Key, incentive conditions: —, goal; ---, control. Groups: G-G, goal unchanged; G-C, goal changed to control; C-C, control unchanged; C-G, control changed to goal.

The post-change period. The introduction of Goal with Encouragement conditions into the Control-Group on the 13th trial produced, as can be seen in Fig. 1, a marked increase in performance level, the group mean score on this trial being in the region of 125 points higher than that of the previous trial under non-incentive conditions. A *t*-test revealed that this increase was a significant one ($t = 3.827$ (D.F. 9) significant at the 0.01 level) and unparalleled in the other three groups. The level thus achieved was maintained throughout the post-change period, improvement being slight but positive. The average increment per trial for the group, as represented by the mean for the regression coefficient 'a', was 1.20.

The mean performance level for the Control-Goal group did, in fact, reach a point which was above that of the Goal-Goal group on the 13th trial and remained consistently above

it for the remainder of the experiment. The '*t*' for the difference in group means for these two groups over the eight trials of the post-change period had a value of 14.053 (D.F. 7) which was significant beyond the 0.001 level. Inspection of individual scores indicated, however, that this 'elation' effect obtained for only some individuals, '*t*' calculated on the individual '*b*' coefficients for the two groups being non-significant.

There was little marked effect of discontinuing Goal with Encouragement in the Goal-Control group, although some slight separation from the Goal-Goal group did occur. The difference in group mean score for this period between the Goal-Goal and Goal-Control groups was significant at the 0.001 level, '*t*' being 6.723 (D.F. 7), but the mean level of Goal-Control subjects was in turn significantly higher than that of Control-Control subjects. The value of '*t*' in this case was 8.554 (D.F. 7), significant at the 0.001 level. Inspection of individual scores revealed, however, that some subjects did decline following the removal of the incentive. Even in these cases, however, the decline was gradual rather than abrupt as in the case of the change in the reverse direction.

IV. DISCUSSION AND CONCLUSIONS

Considered in terms of group trends it is apparent that the results reported here do not entirely bear out the predictions of Hull (1951). Although not clearly confirmed, because of individual differences, the marked reaction of imbeciles to the sudden introduction of an incentive supports, rather, Crespi's (1942) demonstration of a similar 'elation' effect. Broadhurst (1958) has recently confirmed some aspects of Crespi's work. It seems likely that this effect was to a large extent, as Crespi himself points out, a function of the level of expectation built up during the days when no reward was received. It must be remembered that the task was a highly monotonous and fatiguing one, even for imbeciles. By the change point, in fact, some dissatisfaction was being expressed by subjects. In the Control-Control group this became more marked from this point onwards and is reflected in the general decline of this group's performance level.

Two further points are of interest. First, it is noticeable that the 'elation' effect occurred in one trial, suggesting that, although they had previously performed at a lower level than subjects who had received an incentive during learning, individuals in the Control-Goal group had presumably acquired the same degree of skill on the task and were capable, given an incentive, of performing immediately at the level associated with that incentive. It is possible, in fact, as pointed out by Deese & Carpenter (1951), that the rate at which acquisition or habit growth occurs is independent of the level of drive. In any event, it confirms the necessity, suggested by many other authors, of distinguishing 'acquisition' from 'performance' in evaluating the role of motivation in learning.

The second point of interest is the relative permanence of the 'elation' effect. It is possible that, had the experiment been carried further, the performance level of Control-Goal subjects would eventually have trailed back to the level of Goal-Goal subjects, as stated by Hull (1951). There was, however, no evidence of this by the end of the experiment and it would probably have needed a great many trials for this to have occurred. It certainly could not be concluded that the effect was of a temporary nature.

The effect of removing incentives confirmed the findings of neither Hull nor Crespi and did not parallel the change in the reverse direction. It confirms, however, an earlier finding by O'Connor & Claridge (1955) that incentives will have a relatively permanent

effect on the performance level of imbeciles. The persistence of the attitude or set established under Goal with Encouragement conditions was illustrated by the practice of many subjects on the Goal-Control group of continuing to set themselves goals and show obvious pleasure when they claimed to have completed the number of rows they had set as their targets. The ability of imbeciles to 'learn' to set themselves goals in this way indicates that many of the present formulations of such concepts as habit, drive, and incentive may, by themselves, be too simple to explain the complex behaviour which is evident even at this level. This conclusion finds confirmation in many current approaches to the problem of motivation, such as those of Hebb (1949) and Harlow (1953).

At the same time the demonstrations that the 'elation' effect may be paralleled in human subjects suggest that it is through modification rather than rejection of current concepts that a more adequate theory of learning and motivation will emerge. The testing, with human subjects, of some of the basic postulates of animal learning theory would certainly seem to be a fruitful approach to this problem.

The authors wish to acknowledge their indebtedness to Dr R. J. Stanley, Physician Superintendent of Monyhull Hall, Birmingham, who made the patients available, and Dr H. C. Gunzburg, who selected the patients. The work reported was carried out by G. S. Claridge in partial fulfilment of the Ph.D. degree at the University of London.

REFERENCES

- AMMONS, R. B. (1954). Knowledge of performance. Survey of literature, some possible applications, and suggested experimentation. W.A.D.C. Technical Report 54-14.
- BOOK, W. F. & NORVELLE, L. (1922). The will to learn: an experimental study of incentives in learning. *Pedagog. Sem.* **29**, 305-62.
- BROADHURST, P. L. (1958). A 'Crespi Effect' in the analysis of emotionality as a drive in rats. *Brit. J. Psychol.* **49**, 56-8.
- CLARIDGE, G. S. (1956). Factors affecting the motivation and performance of imbeciles. Ph.D. Thesis, University of London Library.
- CRESPI, L. P. (1942). Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.* **55**, 467-517.
- CRESPI, L. P. (1944). Amount of reinforcement and level of performance. *Psychol. Rev.* **51**, 341-57.
- DEESE, J. & CARPENTER, J. A. (1951). Drive level and reinforcement. *J. Exp. Psychol.* **42**, 236-8.
- GOODENOUGH, F. L. (1926). *Measurement of Intelligence by Drawings*. New York: World Book Co.
- GORDON, S. (1953). Some effects of incentives on the behaviour of imbeciles. Ph.D. Thesis, University of London.
- HARLOW, H. F. (1953). Motivation as a factor in the acquisition of new responses. In: *Current Theory and Research in Motivation: a Symposium*. Lincoln, Nebr.: Nebraska University Press.
- HEBB, D. O. (1949). *The Organization of Behavior. A Neuropsychological Theory*. New York: Wiley.
- HULL, C. L. (1951). *A Behavior System*. New Haven: Yale University Press.
- MACE, C. A. (1935). Incentives: some experimental studies. *Rep. Med. Res. Coun. Industr. Hlth Res. Bd, Lond.* London: H.M.S.O.
- MACKWORTH, N. H. (1950). Researches on the measurement of human performance. *Spec. Rep. Ser. Med. Res. Coun., Lond.*, no. 268. London: H.M.S.O.
- O'CONNOR, N. & CLARIDGE, G. S. (1955). The effect of goal-setting and encouragement on the performance of imbecile men. *Quart. J. Exp. Psychol.* **7**, 37-45.
- ZEAMAN, D. (1949). Response latency as a function of the amount of reinforcement. *J. Exp. Psychol.* **39**, 466-83.

(Manuscript received 28 December 1956)

THE EFFECTS OF INCENTIVES ON THE PERFORMANCE OF DEFECTIVE IMBECILES

BY D. WALTON

Rainhill Hospital, near Liverpool and Winwick Hospital, Warrington

AND T. L. BEGG

Public Health Department, Liverpool

The present experiment was designed to inquire into the effects of different incentive conditions on the performance of male imbeciles. In particular the aim was to clarify further the effect of goal-striving and encouragement on performance in uncontaminated experimental groups. The experiment was suggested, first, because the previous results in the field had been equivocal, and secondly, since inquiry into the trainability of imbeciles had been encouraging. The latter results suggested how imperative it was to determine the best incentive conditions to be employed in any experimental workshop in which such work was being undertaken.

Forty-eight imbeciles were used. On the basis of their first performance on the Leg Persistence Test they were divided into four matched groups of 12 subjects each and asked to perform under different motivating conditions. There were four incentive conditions used—'Control', 'Goal', 'Competition' and 'Encouragement'. The experiment continued at the rate of one trial per day over twenty trials. As far as possible each of the four groups was tested on the same day for the equivalent trial number. The four groups were also found to be closely matched on the basis of intelligence.

There can be little doubt that incentives improve the performance of imbeciles on a dull routine task. In contrast, under conditions of indifference their performance deteriorates. No support was gained for Gordon's contention that imbeciles respond to goal conditions of motivation better than when encouraged. In the present study competition and encouragement appear to be the best conditions of motivation.

I. INTRODUCTION

With the publication of Gordon's work (1953) has come a revival of interest in the relatively unexplored field of how low-grade defectives respond to incentives. Although previously Aldrich & Doll (1931) and Gardner (1945) carried out experiments with imbeciles in which incentives were used, it is fair to say that this area of study has been relatively neglected. Such a pessimistic view of the abilities and trainability of imbeciles has been adopted by such authorities as Tredgold (1952) and Lewis (1929) that the study of incentives seemed hardly worth investigating, and somewhat irrelevant. Tredgold (1952), for example, states:

As a result of training, a considerable proportion can be employed in such simple duties as sweeping and scrubbing floors; polishing brasses; weeding garden paths; collecting potatoes; helping in the laundry and so on. But they can only do these things under supervision. They are quite incapable of adapting themselves to anything out of the routine to which they have been accustomed, and it is doubtful whether he work they do from an economic standpoint is worth the supervision it entails. None of them can contribute at all appreciably toward the cost of their keep....

Again, Lewis (1929), writes:

Imbeciles are incapable not only of earning an independent livelihood, but even of contributing materially to their own support. The best that can be expected of imbeciles will be the simplest of routine tasks under supervision. The brightest... can manage in a somewhat irregular fashion, such menial duties as weeping, dusting; scrubbing floors; washing earthenware and unbreakable articles, and even rough laundry work; even here, however, they would need almost continuous supervision.... In wood-work the best of the imbecile youths can manage rough-polishing, sand papering and hammering large nails, and in needle-work the best of the girls can manage coarse stitching and tacking; few will be able to thread their own needles... (and)... tying shoe laces often remains entirely beyond their powers.

It is only recently, however, that the social and psychological aspects of mental deficiency have come to be studied with the development of a systematic psychology of the defective as its principal aim. Amongst these studies have been investigations into the ability of imbeciles to learn skills and be employed doing simple contractual work in sheltered workshops (Tizard & Loos, 1954; Loos & Tizard, 1955; Clarke & Hermelin, 1955). The results from these studies throw doubt on traditional views about the ability and trainability of defective imbeciles. As Clarke & Hermelin (1955) remark: 'It seems that the limits to the trainability of imbeciles are very much higher than have been accepted traditionally either in theory or in practice.'

These early experimental results also suggest the importance of submitting to systematic inquiry, not only the abilities and trainability of defectives, but also the study of the best incentive conditions to be used for this type of patient in such experimental workshops. Contrary to the suggestions of Lewis (1948) and Kallman (1951) that imbeciles constitute a separate clinical entity and that the basic determinants of behaviour found in normals and even in high-grade feeble-minded patients do not apply to them, the work of Gordon (1953), Gordon, O'Connor and Tizard (1954, 1955), O'Connor & Claridge (1955), Tizard & Loos (1954), Loos & Tizard (1955) and Walton & Begg (1955) would suggest that there is no real difference between normals and imbeciles in their response to incentive conditions.

Gordon *et al.* (1954) found, for example, that imbeciles were capable of responding to incentives, and that when they were working toward a goal they did better than when simply encouraged. Both incentive groups did significantly better than the control group who were given the original instructions and nothing else. They concluded that in their manner of responding to incentive, imbeciles did not appear to differ from normals (Mace, 1935). Certain criticisms, however, could be made about Gordon's work. First, very small numbers of subjects were used, particularly in the later experiments. Secondly, the superiority of the 'Goal' group may have depended on Gordon's personal approach to the imbeciles striving towards a goal and their successes or failures rather than 'their own self-competition as an internalized drive'.

O'Connor & Claridge (1955) attempted to make a more clear-cut distinction between 'Encouragement' and 'Goal-striving'. For this purpose they used 'Goal with Encouragement' and 'Goal with Indifference'. Their results indicated that goal-striving by itself did not increase the performance level significantly unless it was accompanied by 'social approval of work' and encouragement.

Loos & Tizard (1955) asked the question: 'What conditions will enable them (i.e. defectives) to work steadily and happily throughout a normal working day without compulsion or punishment?' In the early stages of their investigation they found general confirmation of Gordon's findings in that when imbeciles are given goals to work for there is an improvement in their performance. In one important respect they did not find confirmation of his work. They found, for example, that after two weeks there was a gradual fall. Their conclusion was, 'that in a workshop our subjects' desire to beat their own best performance was not a strong enough motivational force to keep them working steadily throughout a normal working day—after the novelty of the change in conditions had apparently worn off'.

When the incentive conditions were changed in such a manner as to give a good reason why the defective should work fairly consistently there was improvement. In their third experiment the most favourable conditions were those in which the imbeciles worked as

members of small groups where their own production was related to high-grade patients, with whom they worked on a basis of equality. It suggested that the factor of self-esteem was of prime importance in both determining and maintaining an improvement in performance.

The present experiment was designed to clarify further the effects of different incentives on the performance of imbeciles and to determine more accurately the respective roles of goal-striving and encouragement in uncontaminated experimental groups.

II. METHOD

Forty-eight male imbeciles were used. Each subject was individually given the Leg Persistence Test (Gordon *et al.* 1954). First, the procedure was demonstrated to the subject. The subject was then seated with a chair opposite him and given the following instructions: 'When I tell you, I want you to keep your leg above this chair for as long as you can. Do not touch the chair with your leg but try to hold it up for as long as you can without touching the chair. Are you ready? Go!'

The subject's score was the number of seconds he held his leg up in the air.

On the basis of the scores from the first trial, the 48 subjects were divided into four matched groups of 12 subjects each, and asked to perform under different motivating conditions. Four incentive conditions were used: (1) *Control*. The subject was given the above instructions. No encouragement or further comment was made by the examiner. (2) *Goal*. The subject was able to follow his progress by watching a pointer record his persistence. Another pointer showed his level of achievement on the previous trial. As under control conditions, no encouragement or comment was made by the examiner beyond exactly the same instructions as given to the control group. (3) *Competition*. This group watched each other's progress, knowing their own results and the results of the other subjects in the group. As under 'Goal' conditions the 'pointer' was used. One 'pointer' showed the highest score on that day by a member of the group whilst the second recorded the persistence of the remaining subjects. No encouragement or comment was made by the examiner beyond the same instructions given to the control group and an explanation of what was demanded of them. (4) *Encouragement*. No 'pointers' or observable goals were available to the imbeciles in this group. They were given the same instructions as the control group, followed by continued personal contact and praise of the subject by the examiner during each of the experimental trials.

The experiment continued at the rate of one trial per day over a further twenty trials. Except for certain unavoidable delays each of the four groups was tested on the same day for the equivalent trial number. In all, there were 960 observations.

There was no significant difference between the four groups in their initial scores on the Leg Persistence Test. The means in seconds were respectively (Control, $\bar{x}=70.0$; Goal, $\bar{x}=66.2$; Competition, $\bar{x}=69.6$; Encouragement, $\bar{x}=69.1$). The mean Terman-Merrill I.Q.'s and s.d.'s for the Control, Goal, Competition and Encouragement groups were 47.54, s.d. 10.0; 47.90, s.d. 6.4; 45.0, s.d. 8.5; 46.58, s.d. 8.8 respectively.

III. RESULTS

There are several difficulties about the present data. If the experiment consisted of three treatments being applied to different groups of individuals, so that the differences in response obtained between the several individuals within each of the groups (including the fourth control group) all behaved like experimental errors (i.e. quite randomly and independently of the treatment applied), one could proceed to test the effects of the treatments by a simple analysis of variance. The trouble with the present data is that they do not seem to behave in this way.

In the first place, there are clearly as great differences in the variability from trial to trial within each of the groups as there are in their average levels. In fact, the differences between each of the four groups in their variances or standard deviations are very much greater proportionately than the differences in the four means. This must suggest that there is some sort of interaction between the factors being tested and the successive trials: i.e. that

the effect of the three factors on the response obtained cannot be assumed to be the same at each successive trial.

This can be seen from another point of view. The result obtained at each trial, at least in the control and competition groups, seems to be correlated with the result at the preceding trial. The serial correlations present in each of the four groups are given in Table 2. There is no evidence of second or higher order serial correlations except in the control group. All the figures are the coefficients squared, so the square root is needed if the actual correlations are wanted. There is, therefore, evidence of trend or cyclical fluctuation in the results obtained, especially in the control and competition groups. This can be seen quite clearly from Fig. 1, especially at the beginning of the experiment. Of course, one cannot treat 'competition' as if it were a *constant* factor influencing the results obtained, since it must to some extent be variable from trial to trial.

Some may argue, therefore, that it is not appropriate to use the analysis of variance, since the conditions required to make valid tests of significance do not seem to be met by the present data. This is not in the least serious since the obvious way to handle the data is to have a good look at how it behaves, especially by plotting the results graphically, and to see how one can interpret the results best, using common sense and the specialized knowledge one possesses. Significance tests are not, and never have been, a substitute for inspecting the data thoroughly oneself to see what conclusions can be drawn. Often they are either redundant, because the differences are patently significant merely from looking at the data, or they cannot decide doubtful cases because there is some doubt whether the data really meet the requirements and assumptions underlying the tests.

In the present case, the differences between the means of the four groups are obviously quite significant. Indeed, if one carries out an analysis of variance irrespective of the above objections, the *F* ratio works out at 83.89 which (assuming the validity of the tests) is highly significant. The differences between trials do not seem to be significant. The only thing in doubt is whether the difference between the competition and encouragement group is significant. A *t*-test ($t' = 1.8698$, $0.10 > P > 0.05$) suggests it is not, but the *t*-test is open to the same objections as the analysis of variance.

From an inspection of the data several trends are suggested. First, the experiment lasted one month and bears comparison with the work of Loos & Tizard (1955). Both studies reflect the superiority of incentive conditions, but both studies also agree in that when imbeciles are set relatively routine tasks, even under incentive conditions, their performance does deteriorate, though it never reaches such a low level as under control conditions. The results from the control group are of particular significance in relation to what was said at the beginning of the paper. Under conditions of indifference their performance declines. In the present study this started immediately. The work of Tizard (1953) is also of relevance. He found that supervision with indifference produced poor results compared with supervision involving encouragement. Early in the present experiment subjects in the control group complained a great deal about the boring nature of the task. This also spread to the other groups at a later point in the experiment. It is important, however, to make a special point that this group of imbeciles, though of minimal intellectual endowment, appeared to become very disgruntled at the prospect of an uninteresting task before them. This appears to cut right across current theory. Of particular interest was the distractibility of the subjects in the competition group. Some wandered away and started playing with a bat and ball in a far corner of the room.

No support was found for Gordon's suggestion that imbeciles respond to goal conditions of motivation better than when encouraged. A previously stated criticism of his work—that he may have inadvertently encouraged the subjects and thereby have contaminated his goal-group results—appears justified by the relative superiority of the 'Encouragement' group results in the present experiment. The implication of the study carried out by Loos & Tizard (1955), that the factor of self-esteem was of great importance in determining and maintaining an improvement in performance, was however confirmed. The upshot of the results is that when imbeciles are given an uninteresting task to perform their level of achievement can be improved significantly by the use of incentive conditions. In the

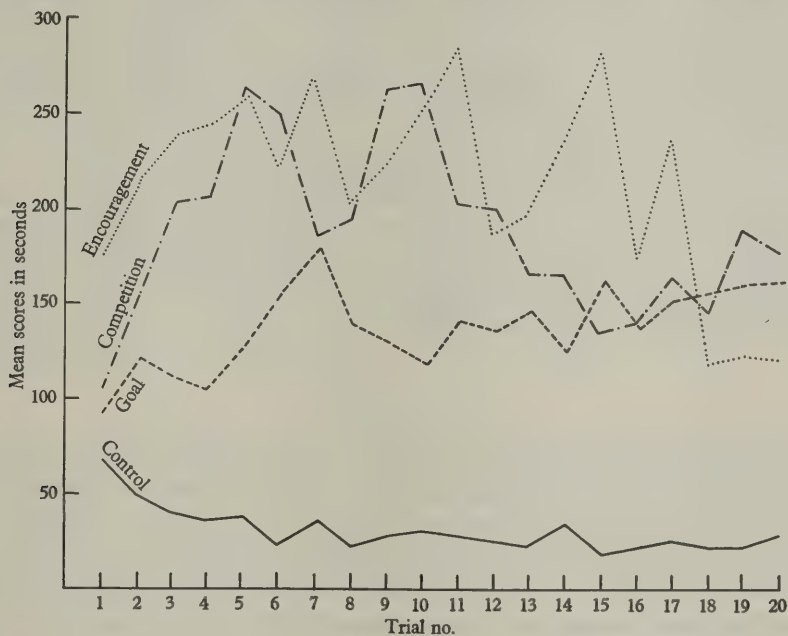


Fig. 1

Table 1. *Group mean scores (in seconds) under experimental and control conditions*

Trial	Control	Goal	Competition	Encouragement
1	68.8	94.8	105.0	176.0
2	50.2	121.0	154.0	219.0
3	44.7	114.0	204.0	239.0
4	35.7	106.0	208.5	244.8
5	36.2	128.7	266.0	261.3
6	23.5	159.5	250.0	220.2
7	36.0	182.1	185.0	268.1
8	23.1	142.0	190.0	202.0
9	30.0	132.5	263.0	221.0
10	32.1	118.0	265.0	250.0
11	30.7	143.0	201.5	283.0
12	27.0	137.9	199.5	177.0
13	26.7	146.3	168.0	192.0
14	35.1	124.8	168.0	230.0
15	18.5	160.8	131.0	281.0
16	21.0	137.5	135.0	175.0
17	26.8	150.0	166.5	238.0
18	24.0	152.0	146.7	117.5
19	24.0	156.8	189.0	123.5
20	29.5	160.7	176.0	121.8

present experiment, 'goal', 'competition' and 'encouragement' conditions of motivation produced significantly better results than under control conditions. The level of achievement under control conditions undergoes a progressive deterioration. Competition and encouragement seem to provide superior motivation conditions than does goal-striving.

Table 2. *Means, standard deviations, variance and serial correlations for experimental and control conditions*

	Control	Goal	Competition	Encouragement
Mean	32.18	138.42	188.58	212.01
S.D.	11.35	20.84	44.56	49.56
Variance	128.89	434.42	1984.74	2455.72
1st order serial	0.474213	0.235826	0.383989	0.1160805
Correlation				
2nd order	0.465343	0.053813	0.048011	0.068185
3rd order	0.227132	0.022548	0.008888	0.020756
4th order	0.145306	—	—	—

Table 3. *Analysis of variance table*

Variance	D.F.	Sum of squares	Mean squares	F ratio	Level of significance
Between groups	3	382,845.18	127,615.06	83.89	0.001
Between Trials	19	28,900.62	1,521.085	1.218	Not significant
Interactions	57	71,174.87	1,248.682		
Trials and interactions	76	100,075.49	1,316.783	—	—
Total	79	482,920.67	61,129.20	—	—

IV. IMPLICATIONS

It has been suggested elsewhere (Lewis, 1948; Kallman, 1951) that imbeciles constitute a separate clinical entity and that the basic determinants of behaviour found in normals do not apply to them. From a comparison of Mace's work (1935), using normal subjects, and the results of the present study plus the work of Gordon (1953), Gordon *et al.* (1954, 1955), O'Connor & Claridge (1955), Tizard & Loos (1954) and Loos & Tizard (1955), all using imbeciles, there appears little or no support for the latter contention. Self-esteem seems to be one factor common to normal and imbecile subjects in determining level of behaviour.

Not only do some of the basic determinants of behaviour appear to suggest themselves as the same in imbeciles as in normals, but recent work is tending to show also that the potential level of achievement in imbeciles is considerably higher than suggested by such authorities as Lewis (1929) and Tredgold (1952). The work of Tizard & Loos (1954) and Clarke & Hermelin (1955) is of particular relevance. The latter authors have shown, for example, that imbeciles can work for long periods and contribute significantly to their keep. Six imbeciles were employed for two and a half years in an experimental workshop as cardboard-box folders. Their average weekly output was 30,000–40,000 for a 35-hour week. Others were able to earn £2 to £3 per week cutting insulated wire to exact lengths using a simple guillotine, others were taught 'soldering of four different coloured wires to the correct terminals of an 8-pin television plug. This requires not only soldering ability but also the distinction of colours and spatial relationship.' Both studies also show that face-value impressions or a subject's initial scores are likely to give a very poor idea of his ability to do tasks after practice. As Clarke & Hermelin (1955) say: 'It seems that to take the imbecile at "face-value" is to ignore his potentialities. These would enable many, after

training, to perform certain useful tasks and to contribute in many cases substantially to their own support.'

The third fact, which has been noted not only in the present study, but elsewhere, is that when imbeciles are set routine, uninspiring tasks to do without incentives, or are left to their own devices without previous training, their performance deteriorates. Clarke tells of a weekend when, in the absence of the supervisor, all six imbeciles from the workshop became bored and gained entrance into the closed workshop and worked all morning without supervision. If one imagines, therefore, a group of hospitalized imbeciles allowed and expected only to perform tasks similar to those suggested by Tredgold and Lewis, then it is not surprising to find that face-value impressions appear correct. As mentioned above, however, controlled experiments have shown that the initial level of performance gave no idea of the ultimate level to which the defective may rise.

If, therefore, imbeciles have been seen to respond favourably to incentives, that they are capable of acquiring 'skills' considerably in excess of those previously designated to them, it seems indefensible to allow them to be simply occupied—'... a more positive approach is needed here quite as much as in other spheres of remedial treatment. Just as, in many countries, rehabilitation centres and sheltered workshops have been established for people with severe physical handicap whose condition was formerly considered hopeless, so it should be possible to aid many of the mentally handicapped who can never be socially independent but who with similar help could contribute substantially to their own support (Clarke & Hermelin, 1955).'

REFERENCES

- ALDRICH, C. G. & DOLL, E. A. (1931). Problem solving among idiots. *J. Comp. Psychol.* **12**, 137-169.
- CLARKE, A. D. B. & HERMELIN, B. P. (1955). Adult imbeciles. Their abilities and trainability. *Lancet*, Vol. 2, 337-9.
- GARDNER, L. P. (1945). The learning of low-grade aments. *Amer. J. Ment. Defic.* **50**, 59-80.
- GORDON, S. (1953). Some effects of incentives on the behaviour of imbeciles. Unpublished Ph.D. thesis, University of London.
- GORDON, S., O'CONNOR, N. & TIZARD, J. (1954). Some effects of incentives on the performance of imbeciles. *Brit. J. Psychol.* **45**, 277-87.
- GORDON, S., O'CONNOR, N. & TIZARD, J. (1955). Some effects of incentives on the performance of imbeciles on a repetitive task. *Amer. J. Ment. Defic.* **60**, 371-7.
- KALLMAN, F. J. (1951). Recent progress in relation to the genetic aspects of mental deficiency. *Amer. J. Ment. Defic.* **56**, 375-81.
- LEWIS, E. O. (1929). Report of an investigation into the incidence of mental defect in six areas, 1925-1927. *Report of the Mental Deficiency Committee*, Part IV. London: H.M.S.O.
- LEWIS, E. O. (1948). Mental deficiency and social medicine. *J. Ment. Sci.* **94**, 258-65.
- LOOS, F. M. & TIZARD, J. (1955). The employment of adult imbeciles in a hospital workshop. *Amer. J. Ment. Defic.* **59**, 395-403.
- MACE, C. A. (1935). *Incentives: Some Experimental Studies*. London: H.M.S.O.
- O'CONNOR, N. & CLARIDGE, G. S. (1955). The effect of goal-setting and encouragement on the performance of imbecile men. *Quart. J. Exp. Psychol.* **7**, 37-45.
- TIZARD, J. (1953). The effects of different types of supervision on the behaviour of mental defectives in a sheltered workshop. *Amer. J. Ment. Defic.* **58**, 143-61.
- TIZARD, H. & LOOS, F. M. (1954). The learning of a spatial relations test by adult imbeciles. *Amer. J. Ment. Defic.* **59**, 85-90.
- TREDGOLD, A. F. (1952). *A Textbook of Mental Deficiency*. London: Baillière, Tindall and Cox.
- WALTON, D. & BEGG, T. L. (1955). Adult imbeciles. *Lancet*, Vol. 2, 616-17.

(Manuscript received 28 December 1956)

A 'CRESPI EFFECT' IN THE ANALYSIS OF EMOTIONALITY AS A DRIVE IN RATS

By P. L. BROADHURST

University of London, Institute of Psychiatry (Maudsley Hospital)

In a previous paper (Broadhurst, 1957), it was shown that, when air deprivation is used as a drive, emotional rats swim significantly faster than non-emotional ones. This effect, which had been predicted by analogy with Eysenck's formulation of neuroticism as an autonomic drive (1955, 1956), was interpreted as reflecting a stronger drive in the emotionally reactive animals. It was not however possible, in the situation used, to decide between two of the possible explanations of the effect observed—a generally higher level of drive, or a greater tendency to drive-arousal on the part of the emotional animals. A simpler experimental situation in which the level of the imposed drive is varied, and which has given rise to the phenomenon termed the 'Crespi effect' was therefore used. Crespi (1942, 1944) showed that when rats who had been trained to run to a small food incentive (low drive) were suddenly rewarded with a larger one, they increased their speed of running to a point beyond that reached by subjects who had been trained from the start under the high drive (large incentive) condition. This was called the 'elation' effect; the converse 'overshoot' in the opposite direction in response to a sudden reduction in drive was termed the 'depression' effect. Crespi's interpretation in terms of emotional drives makes this situation a fitting one for the evaluation of differences due to emotionality. Zeaman (1949) used latency of response in a similar situation, and found both of the Crespi contrast effects when the amount of the food reward was suddenly changed. The depression effect did not reach significance, however. Deese & Carpenter (1951), in a similar study of changes in the latency of response to alterations in drive strength, failed to provide controls by which the Crespi contrast effects could be evaluated. It is, however, clear from their data that the depression effect is absent. The wide generality of the elation effect is shown by the recent finding by O'Connor & Claridge (1958) that imbeciles working at a repetitive task show a significant elation effect in response to a sudden increase in incentive in the form of encouragement and goal setting. These workers also failed to find the depression effect.

The elation effect only will therefore be used in the present experiment, the subjects for which were 48 albino rats, selected from the second and third generations of a selective breeding study of emotionality (Broadhurst, 1958*b*). They had been reared under standard conditions, before their emotional reactivity was observed in Hall's open-field test (1934), under the standardized conditions described elsewhere (Broadhurst, 1958*a*), when they were 104.9 days of age ($\pm$ S.E. 0.35). By virtue of their defecation scores in this test they were assigned to an emotional group (mean score 4.23 ± 0.55 fecal boluses per day) or to a non-emotional group (mean 0.61 ± 0.54). This difference in mean scores is significant beyond the 0.1% level by *t*-test. Subsequent to this testing, but immediately prior to the present experiment, they had served as subjects for a learning task in a Y-shaped underwater discrimination unit, for which four different air deprivation drive levels had been established and used for 15 consecutive days, by submerging the subjects for 0, 2, 4, or

8 sec. under water before release for each learning trial (Broadhurst, 1957). There were thus 12 subjects, six emotionals and six non-emotionals, habituated to swimming to escape from underwater at each of four different levels of drive.

In the present experiment a 10 ft. long straight tank, 6 in. wide, was used. The water-level was 9 in., its temperature was maintained at 20° C. Air temperature averaged 20.8 (± 0.40)° C. A submersible starting cage with a guillotine door which could be opened from above the surface occupied the first 1 ft. of the length of the tank, a wire-netting ramp rising at an easy angle from below the surface occupied the last 1 ft. The intervening 8 ft. were roofed in with stainless steel wire gauze $\frac{1}{2}$ in. below the surface. The subjects were thus forced to swim 8 ft. underwater in order to escape. The time from the release of the door of the starting cage until the subject's snout broke surface at the other end of the tank was recorded to the nearest $\frac{1}{10}$ sec. by stop-watch. These times directly represent speed of swimming since the distance swum is constant. With the same four air deprivation drive levels as previously used, one day's preliminary training was given in this long tank by gradually lengthening the underwater 'ceiling' on successive trials, and on the next day the experiment began with 10 trials of swimming the whole 8 ft. The inter-trial interval was held constant by running those subjects subjected to short delay times in larger groups than those subject to longer ones. On the second day of the experiment the motivation was increased by subjecting all animals to the same delay as that imposed on the two 8 sec. groups, whose motivation was unchanged, and who consequently provided control groups. All animals were given a further ten trials on the next day under the same conditions, and half of each group, randomly selected, another ten on the fourth day. These latter results are shown in Fig. 1, but not included in the subsequent analyses.

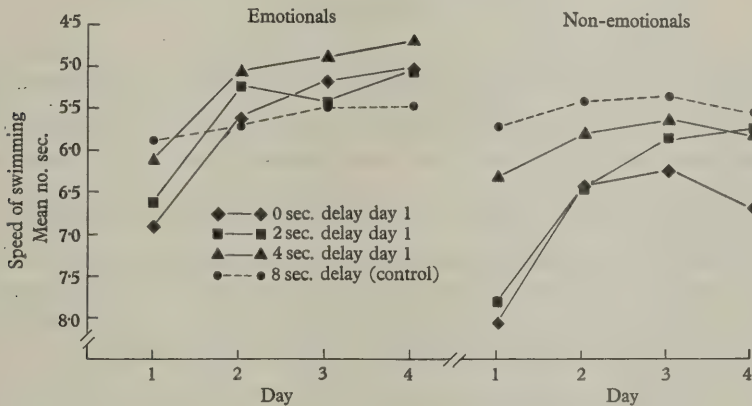


Fig. 1

It is clear from Fig. 1 that the Crespiation effect is present, but among the emotionally reactive rats only. There is some suggestion that the amount of performance 'overshoot' is a function of the magnitude of the change in drive level. The effect appears to persist, or even increase a little, especially among the emotional rats. In order to investigate inter-group differences in level of drive, a 2×4 analysis of variance of the average times each animal returned on the day before the increase in drive (day 1) was performed, after it had been shown that the variance of these data was satisfactorily homogeneous (Pearson & Hartley, 1956). This analysis showed that the F ratios for the two main effects of emotionality and motivation level were significant, with no significant interaction between them. The effect of the four different levels of drive is highly significant, as might be expected ($P < 0.001$); the emotional rats swam significantly faster than the non-emotionals ($P < 0.05$). While this is a relatively lenient criterion, the whole analysis so precisely confirms the findings previously reported (Broadhurst, 1957) that it may be

concluded that emotional rats show a significantly higher drive level than non-emotional ones, as evidenced by their respective speeds of swimming.

In order to investigate the arousal of drive in response to the sudden increase in motivation among the groups previously delayed for 0, 2, or 4 sec., a score reflecting the increase in speed (decrease in time) was computed by subtracting each animal's average time on the second day from that for the first. Each score so obtained was corrected by deducting the average decrease for the six rats in the appropriate control group, similarly calculated. This correction was necessary because the control groups (8 sec. delay throughout) were manifestly also improving their speed (see Fig. 1). A second analysis of variance, like the first, was then performed on these data and revealed no significant effect whatsoever. There was a tendency for the emotionals to increase their swimming speed more than the non-emotionals (mean decrease in swimming time 1.09 and 0.89 sec. respectively), just as there was a consistent trend for a larger increase the lower the original motivation, but neither reached significance. We may therefore further conclude that, within the limits of this experiment, emotionally reactive rats do not display a significantly greater arousal of drive, as measured by an increase in their speed of swimming, than do non-reactive ones, when both groups are subjected to an increase of motivation.

REFERENCES

- BROADHURST, P. L. (1957). Emotionality and the Yerkes-Dodson Law. *J. Exp. Psychol.* **54**, 345-52.
- BROADHURST, P. L. (1958*a*). Determinants of emotionality in the rat. II. Antecedent factors. *Brit. J. Psychol.* **49**, 12-20.
- BROADHURST, P. L. (1958*b*). Studies in psychogenetics: the quantitative inheritance of emotionality in the rat investigated by selective and cross-breeding. *Bull. Brit. Psychol. Soc.* **34**, 3A. (Abstract.)
- CRESPI, L. P. (1942). Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.* **55**, 467-517.
- CRESPI, L. P. (1944). Amount of reinforcement and level of performance. *Psychol. Rev.* **51**, 341-57.
- DEESE, J. & CARPENTER, J. A. (1951). Drive level and reinforcement. *J. Exp. Psychol.* **42**, 236-8.
- EYSENCK, H. J. (1955). A dynamic theory of hysteria and dysthymia. *J. Ment. Sci.* **101**, 28-51.
- EYSENCK, H. J. (1956). Reminiscence, drive and personality theory. *J. Abnorm. Soc. Psychol.* **53**, 328-33.
- HALL, C. S. (1934). Emotional behavior in the rat. I. Defecation and urination as measures of individual differences in emotionality. *J. Comp. Psychol.* **18**, 385-403.
- O'CONNOR, N. & CLARIDGE, G. S. (1958). A 'Crespi effect' in male imbeciles. *Brit. J. Psychol.* **49**, 42-48.
- PEARSON, E. S. & HARTLEY, H. O. (1956). *Biometrika Tables for Statisticians*, vol. 1. Cambridge University Press.
- ZEAMAN, D. (1949). Response latency as a function of the amount of reinforcement. *J. Exp. Psychol.* **39**, 466-83.

(Manuscript received 24 January 1957)

A PATTERN RECOGNITION METHOD OF LEARNING MORSE CODE

BY M. D. ALLAN

Senior Psychologist at the Branch of Scientific Advisor to Air Ministry

This study is concerned with a comparison between the traditional 'analytic' method of learning the morse code, and a 'Gestalt' approach called the Pattern Recognition method. The Pattern Recognition technique consists of transmitting characters right from the start at a speed corresponding to 20 words per minute. There is no preliminary memorizing of the alphabet. All the pupils have to do is to listen to the whole character (i.e. a letter or a number) transmitted at 20 words per minute and, when the instructor has written its meaning on the blackboard, to copy it down. This sort of drill is continued until all the 36 characters (26 letters and 10 numerals) are instantly recognized by their distinctive sound patterns. The pilot experiment showed that pupils trained by the Pattern Recognition method had a significantly better knowledge of the alphabet, after the same amount of training, than pupils trained analytically. As an applied technique in a short intensive course, the Pattern Recognition method produced significantly fewer unsatisfactory pupils during the training period. In a long course, where pupils practised at whatever speed they were capable, pupils trained by the Pattern Recognition method arrived at high speeds earlier than those trained analytically.

I. INTRODUCTION

Morse reception is a speed skill involving the ability to recognize and name auditory patterns. The patterns are composed of different combinations of short and long sounds, each symbolizing a letter or a numeral. In teaching such a skill, a choice must be made between two main varieties of learning theory—the characters can be studied and learned analytically, or they can be treated as 'Gestalten' or indivisible patterns of sound. The oldest method of teaching the morse code, and one still in common use, is an analytic method. The student is first given a card in which all the letters of the alphabet and the ten numerals are represented as visual dots and dashes. He is expected to study this analytically and memorize the parts. The characters are then tapped out slowly by the instructor so that the student can distinguish clearly the 'dits' and 'dahs' which he hears, and match them up with the dots and dashes which he has already memorized. The speed of transmission is increased gradually from day to day until the required speed of reception is attained. This change of speed involves changes in the duration of the elements because everything must be scaled to conform with the international agreement that: (a) a dot is one unit, (b) a dash is equal to three units, (c) the space between component parts of a character equals one unit, and (d) the space between any two characters equals three units. By this method, the proportional relationship between the elements remains constant throughout the learning period, although the duration of the separate long and short sounds changes as the speed of transmission changes. It is difficult to say to what extent this alters the sound of individual characters. Certainly, the auditory impression of a character transmitted in this way at 4 words per minute must be very different from the auditory impression of the same character transmitted at operational speed—say 20 words per minute.

The fact that the morse code is constructed entirely of different combinations of short and long sounds makes the analytic method very attractive—it straightaway suggests

the memorization of components. Yet a morse character is a distinctive pattern of sound—it is a whole sound, and it must be experienced immediately as such before operational speed is possible. The skilled operator has no time to count up the separate elements, so at some point during the learning process he must stop listening to morse characters analytically and instead experience the sound patterns as wholes. This consideration suggests that it might be profitable to forget that morse characters can be so easily analysed, and ask instead if it is possible for beginners to be trained from the start to recognize and name 36 different patterns of sound (i.e. 26 letters and 10 numerals) by experiencing the sound only. This is the ‘Gestalt’ approach, and it demands that individual morse characters should be regarded as indivisible patterns of sound and transmitted at a constant high speed right from the start. Slow speeds are obtained by increasing the spaces between characters.

II. OUTLINE OF PREVIOUS RESEARCH

The idea of learning individual characters at high speed was first suggested in 1918 in a bulletin issued by The Federal Board for Vocational Education (Washington), but appears to have passed unnoticed, probably because there was no supporting evidence. In 1919, Lipmann discovered that an expert operator found it difficult to prolong dots and dashes at low speed to conform with the fixed proportional relationship between elements cited above, and in practice he tended to decrease his speed by increasing the gap between characters. This led Lipmann to suggest (1928) that individual characters should be transmitted at a constant rate of 20 words per minute and slowness for beginners obtained by increasing the spaces between characters. So far, no experimental evidence existed to support these recommendations, but Biegel (1932) and Koch (1936), recognizing that a morse character at high speed has a distinctive rhythmic character, investigated the matter from the point of view of Gestalt theory. These investigators produced some impressive, but inconclusive, experimental evidence in support of the view that a morse character should be treated as an indivisible pattern of sound. The first attempt to compare the relative merits of the ‘Analytic’ and the ‘Gestalt’ approach to morse training was made by Taylor (1943*b*) in a carefully controlled experiment involving 13 subjects covering a training period of 40 hours. His findings showed no significant difference between the two methods in speed of learning.

The generally accepted training in the Royal Air Force has for years been the analytic form and, since no complaint could be made about the standard of skill eventually achieved by operators, probably no change would have been contemplated. The matter, however, came under consideration in the Royal Air Force because of complaints that pupils experience, during the early stages of training, a degree of strain out of proportion to the difficulty of the task. There were many complaints of headaches during training, and the Service records showed a fairly high casualty rate due to apparent lack of aptitude. It occurred to the writer that the strain may be caused by pupils analysing when they ought to be listening. The analytic method demands a summation of elements, and the effort to do that at increasing speeds must impose considerable strain. Furthermore, the fact that the duration of dots and dashes changes with the rate of sending means that the sound of individual characters must change as speed increases. Consequently, during the learning period, continual adaptation must be made to changes in auditory experience,

although such adjustment is not a feature of the learned skill. These considerations suggested that it might be profitable to re-examine the recommendations of Lipmann, Biegel and Koch that speed of transmission of individual characters should be held constant, and that variations in speed should be obtained by varying the interval between characters. Thus, at low speeds, the individual letters or numerals would be transmitted at a constant high speed but the interval between each would be lengthened to produce the over-all low speed. Permission was obtained for the matter to be investigated on Royal Air Force personnel in a live training situation, on the understanding that the normal training schedule would not be interrupted.

III. PLAN OF INVESTIGATION

The plan of investigation consisted first of a pilot study, the purpose of which was to find out if pupils were able to learn morse characters by the method of direct presentation, and if so, to obtain some measure of how effective the learning was, compared with a more orthodox method.

(1) *The Pilot Experiment*

A class of 64 airmen about to start training as wireless operators were taught the morse alphabet by the Pattern Recognition method: that is, individual morse characters were sent out by automatic transmitter at a speed corresponding to 20 words per minute with an interval of 7 seconds between each character. Both the instructor and pupils listened and, as each letter came over, the instructor wrote it on the blackboard, and the pupils copied it down. In writing up the letters, the instructor used 5-letter groups so that the pupils should get used to the form of the 5-letter code. Thus, right from the start, pupils had to listen to whole characters, and they were not permitted to memorize analytically either from a card or from studying visual symbols. The idea was to develop recognition and to establish, by mechanical means, an association between the sound pattern and its meaning which would lead eventually to automatic identification of the sound.

Table 1. *Test of knowledge of letters in 6 words per minute test*

	Pass	Fail	χ^2	P
Pattern Recognition method, $N=66$	51 (77%)	15 (23%)	} 16.1	0.001
Analytic method, $N=71$	31 (44%)	40 (56%)		

To test progress a comparison was made between this group and a control group of 71 men trained by analytic techniques on the same station after the same amount of training. The test was on mixed letter groups at 6 words per minute, which is not really a test of speed but a test of knowledge of the alphabet.

The highly significant result (see Table 1) not only showed that the pupils were learning the code by direct presentation, but that their knowledge of the alphabet was better, after the same amount of training, than that of analytically trained pupils.

(2) *Applied experiments*

The next step was to incorporate this Gestalt approach into a complete training scheme and try it out experimentally against matched control groups trained by the analytic method. This was done at two different R.A.F. units, one on a short intensive course, and the other on a long course where morse instruction is spread out over 18 months.

(a) *Short intensive course.*

The object of the course is to train operators to receive and send morse in plain language, letter code, and letter/figure code at 20 words per minute.

A sample of 41 men were trained throughout by the Pattern Recognition technique and their progress compared with a control group of 39 men trained in the ordinary analytic way. A direct comparison in terms of pass/fail results in progress tests was inconclusive because, in the interests of training economy, some reshuffling of personnel takes place quite early in training. Pupils who are not making satisfactory progress are either suspended or reclassified for further training. A pass/fail comparison was made up to

the speed of 11 words per minute, and this is shown in Table 2. So far as they go, these results indicate that the ability to cope at low speeds is about the same whichever method is used, but that the Pattern Recognition method makes higher speeds easier. A more decisive result is obtainable when the groups are compared in regard to losses sustained through inability to cope with morse. The losses may not be absolute losses, but they do indicate unsatisfactory progress for, in addition to 'Suspensions' they include 'Reclassing' for reasons involving unsatisfactory progress in morse. A comparison in these terms was made covering the whole training period up to the level of operational skill, and it was found that the Pattern Recognition group had significantly fewer losses ($\chi^2=6.15$, $P<0.02$).

Table 2. *Comparison of training progress in standard tests up to 11 words per minute*

Speed	Pattern Recognition method, <i>N</i> = 41		Analytic method, <i>N</i> = 39		<i>P</i>
	Pass	Fail	Pass	Fail	
8 w.p.m.	30 (73%)	11	30 (77%)	9	0.90
9 w.p.m.	29 (71%)	12	23 (59%)	16	0.30
10 w.p.m.	32 (78%)	9	24 (62%)	15	0.10
11 w.p.m.	33 (80%)	8	21 (54%)	18	0.01

Table 3. *Comparison of training progress in terms of practising speeds*

	18 w.p.m. and over	15 w.p.m.	12 w.p.m.
Experimental group (Pattern Recognition), <i>N</i> = 38	14 (37%)	15 (39%)	9 (24%)
Control group (Analytic), <i>N</i> = 44	2 (5%)	23 (52%)	19 (43%)

$$\chi^2 = 13.86; P = 0.001.$$

(b) *Long course.*

Again it was not possible to make comparison of final pass/fail training results because, after such a long training period, very few pupils fail. However, it is the procedure at this particular school to allow pupils to practise at whatever speed they are capable. Here was a basis for comparison, provided that the comparison was made not too near the final examination where differences tend to be levelled out. In Table 3, the two groups are compared at a stage where they are normally expected to be receiving at 12-15 words per minute. The difference has high statistical significance (for 2 degrees of freedom, $\chi^2 = 13.86$, $P = 0.001$) and is further evidence that the Pattern Recognition method makes the learning of morse easier.

(3) *Further observations*

In addition to operative skill, pupils must also qualify in technical knowledge. Some indirect evidence that the Pattern Recognition method involved less strain was obtained when two of the groups were compared for technical knowledge. There was no statistical significance between the groups in terms of their intelligence ratings, yet in the final examination the group which learned morse by the Pattern Recognition method was superior to the control group at the 5% level of significance for 3 degrees of freedom. It is difficult to explain this since both groups had the same sort of instruction and the same standard of examination. It had been hoped that the new method would make learning morse less fatiguing on the operative side, so a possible explanation is that the pupils in the experimental group were less fatigued, and therefore more able to absorb theoretical knowledge.

IV. THE EFFECT ON SENDING

Sending has never presented the same difficulties as receiving, so no special study was made of sending problems other than to see if the Pattern Recognition method of learning to receive would affect adversely learning to send. Normally pupils practise sending at very slow speeds at the same time as they practise receiving at slow speeds. The 'Gestalt' approach, however, rests upon preserving the distinctive wholeness of the characters, so formal practice in sending was postponed until pupils were thoroughly familiar with

receiving characters as wholes. During that time, however, pupils were shown the correct stance and method of holding the key, and were encouraged to fill in odd moments with flexibility exercises and even to try to reproduce with the key, during the 7 seconds interval, the patterns transmitted on the receiving tapes. Since no comparative studies were made on sending, nothing can be reported other than the fact that the method had no noticeable adverse effect on learning to send.

V. DISCUSSION

The results of these experimental field studies point clearly to the usefulness of the Gestalt approach in learning what is, in fact, a recognition skill. The full benefits are more evident in the later stages of training than at the beginning. Perhaps this explains the non-significant results obtained by Taylor when he compared the two methods. His pupils were trained for 40 hours which, in an ordinary training scheme, is about the length of time required to bring the average beginner to a speed of 8 words per minute. There are limits to the extent to which the learning of an automatic skill can be speeded up by any method, and it seems, from the present study, that Taylor may have stopped his experiment too soon. The comparisons which are recorded in this paper began after about 40 hours of practice, and there was found to be little difference between the groups in the early comparisons, but the differences became significant as the skill approached operational level.

As well as making morse easier at higher speeds the Pattern Recognition method puts less strain on the learner. This could only be assessed indirectly as evidenced by the higher performance of the experimental group in technical knowledge (see § III (3) above) but it was very apparent to instructors who described their pupils as being 'less worried than usual'.

The method of instruction under discussion has been described as a 'Gestalt' approach, but it would be wrong to read into this more than the fact that the material is presented to the pupil in the form of whole patterns. Whether or not they react strictly in accordance with Gestalt theory is not at all clear. In the experiments, it was observed that they quickly became aware of the distinctive pattern of the sounds they were listening to; this was evidenced by the sharp reaction of the whole group when a stray character was introduced during one of the early lessons. But when it came to distinguishing one pattern from others and identifying it, there was evidence that many pupils were attending to some detail in relation to the whole pattern. This is a very different process from the memorizing of details which characterized the analytic method. There the details were wholes in their own right and the learning process was an additive one. In the Pattern Recognition method attention to detail, when it was found to occur, was always relative to the whole structure, and was more in the nature of selecting cues which would help to fix the identity of the whole. This brings the study into the realm of perception, for even a synthetic skill such as morse undoubtedly involves auditory perception. Intrinsically there is little to distinguish it from any of the recognition skills of everyday life which are acquired without formal training. These everyday skills, such as recognizing tunes or birds' songs, may involve much knowledge of detail, but the recognition process, where it is an involuntary one, always starts from experiencing the whole.

Investigators in the field of industrial operations have recently been drawing attention

to the perceptual element involved in speed skills and indicating that it is the more difficult element to master (Seymour, 1956). The present study is primarily directed towards solving the perceptual difficulties of morse, and it would appear that the repeated association of whole patterns of sound with their meaning allows free perceptual organization to take place on a subliminal level, rather characteristic of the way the recognition skills of everyday life are learned.

REFERENCES

- ANON. (1918). Emergency war training for radio mechanics and radio operators. *Bulletin Federal Board for Vocational Education*, no. 16. Washington: Government Printing Office.
- BIEGEL, R. A. (1932). Das Anlernen der Höraufnahme durch Funkentelegraphisten. *Psychotech. Z.* **7**, 147-51. (Reviewed by Taylor, 1943, *Psychol. Bull.* **40**, 468.)
- KOCH, L. (1936). Arbeitspsychologische Untersuchung der Tägigkeit bei der Aufnahme von Morzezeichen, zugleich ein neues Anlernverfahren für Funker. *Z. angew. Psychol.* **50**, 1-70. (Reviewed by Taylor, 1943, *Psychol. Bull.* **40**, 469-71.)
- LIPMANN, O. (1919). *Z. angew. Psychol.* **15**, 301-40. (Taylor, 1943, *Psychol. Bull.* **40**, 468.)
- LIPMANN, O. (1928). Eignungsprüfungen für Funker. In Abderholden, E., *Methoden der experimentellen Psychologie*. Berlin: Urban und Schwarzenburg. (Taylor, 1943, *Psychol. Bull.* **40**, 468.)
- SEYMOUR, W. D. (1956). Experiments in the Acquisition of Industrial Skills. *Occup. Psychol.* **30**, 103.
- TAYLOR, D. W. (1943*a*). Learning Telegraphic Code. *Psychol. Bull.* **40**, 461-87.
- TAYLOR, D. W. (1943*b*). The learning of Radiotelegraphic Code. *Amer. J. Psychol.* **56**, 319-53.

(Manuscript received 13 February 1957)

THE DEVELOPMENT OF INDEPENDENT MUSCULAR CONTROL OF SEPARATE EYELIDS IN TWO RACIAL GROUPS, EUROPEAN AND POLYNESIAN

By P. R. N. SUTTON

University of Melbourne

The ability to control separate eyelids independently ('the voluntary winking ability') is studied in 527 subjects of European stock and in 262 Polynesians, with the following results: (1) In both groups, there is a significant correlation between voluntary winking ability and age. (2) Winking ability of the males is not significantly greater than that of the females. (3) In children under 10 years of age, the incidence of non-winking is significantly higher than in the remaining subjects. (4) In the European group, but not in the Polynesian, the incidence of those who could wink both eyes readily and with approximately equal facility is significantly lower in those who are 10 years and under, than in subjects over 10. (5) In those subjects of European stock who could wink one eye only or could wink one eye more easily than the other, a preponderance of the use of the left eye for winking is found. This is mainly due to the preference for winking the left eye which is seen in those who are under 10 years. However, in the Polynesians there is no significant difference between the frequency of winking each eye.

I. INTRODUCTION

Voluntary unilateral winking was considered by Duke-Elder (1952) and by Danielson (1932) to be an act which is usually established early in life and which it is difficult to change. The ability to wink the eyes independently has been observed in 789 subjects, 527 of them being of European stock and 262 of Polynesian, the age range in both groups being from 3 years to over 25 years, and some associations have been observed of which no record has been found in the literature. The results obtained in the two racial groups in regard to the ability to wink one eye voluntarily at a time are compared in this paper; and, for each group, the influence of the age and sex of the subjects is considered.

The influence of sex and of age on winking ability. W. Rider (1898) found that in 'older persons. . . there is still a small residuum of non-winkers, eleven out of 1343 men and seventy-three out of 626 women'. He felt that this 'disproportion between the sexes in the matter of winking ability is probably the result of difference in occupations and manner of life. In her occupations, woman finds less use for monocular vision, and, her life being less out-of-door, her eyes do not require lid protection so frequently.' Ten years earlier, C. Rider (1889) had reported that 30 females out of 180 could not wink, but only one male was unable to do so out of 120 examined. Danielson (1932) said 'I believe with the Riders that men wink better than women'. So far as age is concerned, Duke-Elder (1952) was of the opinion that 'the habit' (winking) 'is usually acquired in early childhood'. Danielson (1932) used similar terms, stating 'I believe from my present work that the winking ability is determined in early childhood and persists throughout life much the same as handedness'.

Changes in winking ability. In this regard, Danielson (1932) stated that: 'While I have not had a sufficiently large group for dependable data, I am quite certain that the only condition that will ordinarily ever markedly change or reverse a unilateral winking habit or modify a non-winking condition is a unilateral painful lesion associated with blepharospasm.' Duke-Elder (1952) felt that the habit 'is as difficult to change in later life as

handedness'. Schmidt-Rimpler (1889) said that: 'Some individuals find it difficult to close the one eye, and only learn it by long practice.' However, McKay (1889) cited his own experience of a spontaneous increase in winking ability.

II. PLAN OF INVESTIGATION

The subjects. The 527 subjects of European stock were consecutive patients presenting for dental examination in Melbourne, with the exception of two, each of whom was blind in one eye, and of those who, because of the loss of maxillary incisor teeth, were not included in the main investigation, of which this study is, as it were, a by-product. The age and sex distributions are shown in Table 1A. The Polynesian sample was obtained on the island of Rarotonga, situated in the South Pacific about half-way between Tahiti and

Table 1A. *European stock*

Age in years...	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21+	Total
Male	4	15	9	13	25	13	14	23	12	19	16	11	12	9	8	3	6	2	22	236
Female	2	10	16	17	16	17	19	27	12	17	21	13	11	9	6	7	7	8	56	291
Total	6	25	25	30	41	30	33	50	24	36	37	24	23	18	14	10	13	10	78	527

Samoa. In spite of the remote situation of the island, there has been a certain amount of intermarriage with Europeans, but it is considered that the subjects included in this investigation were almost entirely of Polynesian stock. They comprised all the children attending the primary school of the most remote village on the island, all adults attending for examination at the most remote meeting house, and all the pupils of the secondary school who were 15 years of age or more. Their age and sex distribution are shown in Table 1B.

Table 1B. *Polynesians*

Age in years ...	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21+	Total
Male	3	3	6	13	10	17	6	6	6	4	7	10	12	7	3	1	3	2	12	131
Female	2	1	3	10	14	2	9	6	12	8	7	7	19	7	4	5	5	2	8	131
Total	5	4	9	23	24	19	15	12	18	12	14	17	31	14	7	6	8	4	20	262

Classification of winking ability. In this paper, by winking is meant 'to close one eye, as in aiming at a target' as defined in the *Oxford Dictionary* (1933), and not in the sense, as used by Miles (1931), of the simultaneous closing of both eyes. Six categories of eye winking were differentiated: (1) Subjects who could not close either eye without closing the other one also, in this paper termed non-winkers. (2) Those who could close only the right eye without closing the other, here called unilateral right winkers. (3) Those who could close only the left eye without closing the other, called unilateral left winkers. (4) Those who could close the right eye more easily than they could the left, termed preferred right winkers. (5) Those who could close the left eye more easily than they could the right, termed the preferred left winkers. (6) Those who could close both eyes readily and with approximately equal facility, and called in this paper the bilateral winkers. This classification is similar to that used by W. Rider (1898) in his study of visual acuity. The closing of one eye was, in most cases, accompanied by the puckering of the other. This was disregarded, although its degree varied greatly. In most cases, where difficulty was experienced in closing an eye, a considerable degree of contraction of the accessory facial muscles was seen. Little difficulty was experienced in deciding to which category each

subject should be assigned. In order to assess the error of these observations, 106 European subjects and 52 Polynesians were re-examined several weeks after their initial testing. Only one discrepancy was noted in each group, in both cases the difference being of only one winking category.

III. RESULTS

The results are summarized in Table 2.

Table 2

	Europeans		Polynesians	
	Number	Percentage	Number	Percentage
Non-winkers	51	9.7	12	4.6
Unilateral right	50	9.5	24	9.2
Unilateral left	79	15.0	26	9.9
Preferred right	103	19.5	68	26.0
Preferred left	128	24.3	62	23.7
Bilateral winkers	116	22.0	70	26.6
	527	100.0	262	100.0

Non-winking. In those subjects who were 10 years of age and under, in both racial groups, the incidence of non-winking was significantly higher ($P < 0.001$) than in those over 10 years old.

Bilateral winking. In those who were 10 years of age and under in the European group, the incidence of bilateral winking was significantly lower ($P < 0.01$) than it was in those subjects who were over 10 years old. However, in the Polynesians, too many subjects developed this skill at an early age for the last-mentioned effect to be demonstrable.

Unilateral and preferred winking. Among those subjects who were unilateral or preferred winkers, in the European group 207 (57.5%) winked the left eye and 153 (42.5%) the right one ($P < 0.01$). This preponderance of left winking is mainly due to the winking preference of those who were 10 years of age and under, of whom 100 winked the left eye and 65 the right, this difference in proportion, also, being significant ($P < 0.01$). However, in the Polynesian group, both in the total number of subjects in these classes and in those under 10 years of age, whether the results of the unilateral and of the preferred winkers were added or considered separately, there was no significant difference between the number of subjects who winked the right eye and those closing the left one.

Winking ability and sex. Both in the European and in the Polynesian groups, no significant difference was established between the proportions of the two sexes which were seen in each of the winking categories.

Winking ability and age. In the European group a significant correlation ($P < 0.001$) was found, in each sex, between winking ability and age, the difference between the two sexes again not proving to be significant, the pooled coefficient of correlation being 0.39. The Polynesian group gave a similar result, r being 0.26 ($P < 0.01$). These relationships are displayed in Fig. 1, in the compiling of which the comparative winking ability was expressed by allotting the numerical value of 0 to the non-winking groups, of 1 to the unilateral winking groups, of 1.5 to the preferred winkers, who could wink one eye with ease and the other with more difficulty, and of 2 to the bilateral winking groups. The ages were grouped into 3-year periods to improve the clarity of the graphs. In Fig. 2 the upper curves display the rise in incidence of the ability to wink either eye (bilateral and preferred winkers), and the lower curves set out in Fig. 2 depict the fall in the percentage of non-winkers as age increases.

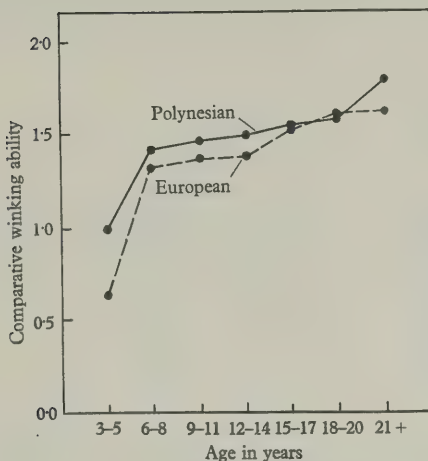


Fig. 1

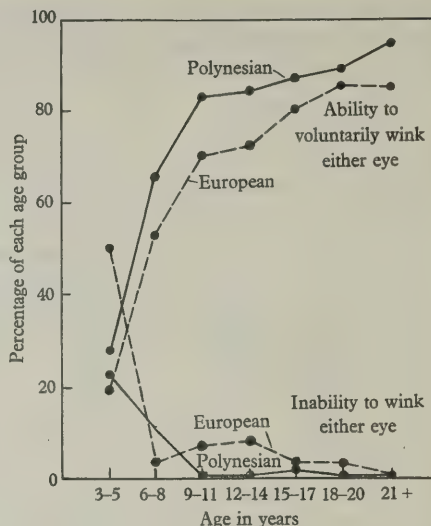


Fig. 2

Fig. 1. The relationship between winking ability and age in years in the European and the Polynesian racial groups.

Fig. 2. The relationships between the ability to voluntarily wink either eye, and age in years; and between the inability to wink either eye independently of the other, and age in years. European and Polynesian racial groups.

IV. DISCUSSION

The incidence of right winking ability in the subjects of European stock is the same as that found by Danielson (1932) in a slightly smaller sample. However, the incidence of non-winking, 9.7%, was higher than the 7% which he reported, and that of bilateral winking (22%) was lower than his figure of 33%. In addition, the occurrence of left winking (39.3%) was higher than his finding, which was 31%. Unfortunately, Danielson did not state the ages of his subjects, but it is considered that the difference between his results and those of the subjects of European stock in the present work may be attributable to there being, in this sample, a larger proportion of subjects who were 10 years of age and less; for, in this group, there is a significantly greater proportion of those who are non-winkers and of those who wink the left eye, and a significantly smaller proportion of bilateral winkers than are to be found in the remainder of the sample.

These Polynesians spend most of their time in the open air, and their work includes few tasks which require monocular vision. Their winking pattern differs from that of the European group mainly in that their winking ability was similar on both sides, and from that group and from the data recorded by Danielson (1932) in that only 4.6% of the Polynesians were unable to wink.

The correlation between voluntary winking and age, as revealed in the present work, is of interest, when the belief of Danielson (1932) is recalled that 'winking ability is determined in early childhood and persists throughout life much the same as handedness': and that he was 'quite certain that the only condition that will ordinarily ever markedly change or reverse a unilateral winking habit or modify a non-winking condition is a unilateral painful lesion associated with blepharospasm'. Duke-Elder (1952) also was of the opinion that 'the habit' (winking) 'is usually acquired in early childhood and is as

difficult to change in later life as handedness'. The relationship between age and winking ability from 6 years to 21 years of age is probably linear, there is certainly a highly significant correlation. When the subjects under 6 years of age are included the curve is, obviously, no longer linear, and it is possible that the youngest age group forms a different population. This is suggested, also, by the similar shape of the curves obtained from the data of the two racial groups. However, the precision of the measurement is not sufficient to allow any definite statement on this point. There is still a highly significant correlation between winking ability and age; in fact, for obvious reasons, r has increased.

By excluding the non-winkers from the data, it can be demonstrated, in both races, that the shift of the point on the curve (Fig. 1) which represents the youngest age group, from the approximately linear conformation of those points which denote the winking ability of the other age groups, is due to the high incidence of non-winking in this group. This is displayed in the lower curves of Fig. 2. In this connexion the concept of C. Rider (1889) may be recalled, that there is an 'age at which monocular vision is sought by the individual . . . I think it will be found to be at the age of about five years in boys, later in girls.' In the present study, in the European group, only 42% of those who were under 5 years old were able to voluntarily wink one or both eyes. However, this ability was seen in 80% of the group of children who were 5 years old, and it was present in over 93% by the time they were between 6 and 7 years of age. In the Polynesians, 55% of those under 5 years could wink, and the figure rose to 87% in those who were 6 years of age.

C. Rider in 1889, W. Rider in 1898, and Danielson in 1932, were all of the opinion that 'men wink better than women' (Danielson, 1932). However, in the present work, both in the Polynesians and in the subjects of European stock, the winking ability of the male group was not significantly superior to that of the female one. W. Rider (1898) said that 'this disproportion between the sexes in the matter of winking ability is probably the result of differences in occupations and manner of life'. This theory is supported by the present results, for when the European figures are compared with those presented in his paper, it is seen that the altered way of life of women during the present century has been accompanied by a significant decrease in the incidence of non-winking in those of them who are over 21 years of age. No conclusion can be drawn in regard to the younger age groups, as no figures regarding the winking ability of these groups in the past have been found.

I should like to express my thanks to Prof. Sir Arthur Amies, C.M.G., Dr J. Ringland Anderson and Dr S. B. Hammond, Senior Lecturer in Psychology, University of Melbourne, for reading the draft of this paper; and to Mr R. T. Leslie, Senior Lecturer in Mathematical Statistics, University of Melbourne, for his advice on statistical matters.

REFERENCES

- DANIELSON, R. W. (1932). Some theoretical and practical points regarding unilateral voluntary winking. *Amer. J. Ophthal.* **15**, 611-17.
- DUKE-ELDER, W. S. (1952). *Text-Book of Ophthalmology*, Vol. v, p. 5153. London: Kimpton.
- MCKAY, R. J. (1889). In discussion on paper by Rider, C. E. (1889).
- MILES, W. R. (1931). Elevation of the eyeballs in winking. *J. Exp. Psychol.* **14**, 311-32.
- RIDER, C. E. (1889). The winking test. *Trans. Amer. Ophthal. Soc.* **5**, 551-59.
- RIDER, W. (1898). Unilateral winking as a test of comparative visual acuity. *Trans. Amer. Ophthal. Soc.* **8**, 434-52.
- SCHMIDT-RIMPLER, H. (1889). *Ophthalmology and Ophthalmoscopy*. Quoted from Rider, C. E. (1889).

(Manuscript received 20 March 1957)

SHAPE RECOGNITION: A REPLY TO DODWELL

By J. A. DEUTSCH

Institute of Experimental Psychology, University of Oxford

Dodwell (1957) criticizes the writer's theory (Deutsch, 1955) on three main grounds. First, he says that Lashley's (1938) experiments cannot be used as evidence for it, since his figures subtended too large a visual angle to be seen as a whole, while the theory postulates that they are seen as a whole. The reason, says Dodwell, why they could not be seen as a whole is that parts of them would fall outside the rat's area of distinct vision. However, Lashley's square would have subtended about 30° and is readily seen as a whole by the human eye, even though the edges be outside the area of distinct vision. We cannot form conclusions about figures being seen as a whole from considerations of distinct vision or acuity.

Secondly, he objects that rats do not treat shapes which have been rotated in the sagittal plane as equivalent. It is true that there is no good evidence for this in the rat, which has a preference for cues arising from the tilt of the figure. However, as there is no good evidence against either, and as other species appear to recognize shape as apart from tilt, it seems sound enough to frame a theory which explains this ability. In any case, the Lashley jumping stand on which the rat is tested may prejudice the issue in favour of tilt. A rat will jump to horizontal bars without training. The selection of cues of tilt has an obvious biological utility in jumping situations.

Thirdly, Dodwell objects to the writer's theory because he has found that a few of his rats can discriminate circles from squares (apparently without any training). Apart from the suspicious nature of these particular findings, Lashley did find that this discrimination presented the utmost difficulty, and Dodwell has not explained in a satisfactory manner why this should have been the case. He suggests that the brightness relationship in the various quadrants of a card containing a solid circle and a solid square are the same. Therefore, as brightness is a 'dominant' cue, the rat fails to discriminate. This line of reasoning would lead us to predict that a rat could not discriminate a square from another rotated through 45° . However, Lashley found that rats could perform this discrimination.

In connexion with the second criticism, Dodwell quotes Gellermann (1933) to show that chimpanzees do not have the ability to recognize rotated triangles, as demanded by the writer's theory. This is because they rotated their heads in the discrimination situation. But it seems that this is a misinterpretation of the evidence. It is clear from Gellermann's account that head rotation took place after recognition of the correct figure, and not before. Head turning only occurred in relation to the correct figure, even when new figures were substituted for the negative figure. 'All subjects continued to react positively to the triangle under these conditions. An interesting type of behaviour accompanied these correct responses. In every case the subjects hesitated on their trials under these new conditions, and, after looking from form to form *they turned their heads about 60 degrees to one side or the other and viewed the triangle*. Then they made the correct response.' (Gellermann's italics.) Such findings support the writer's theory, not Dodwell's.

Dodwell's rebuttal of the writer's theory of shape recognition consists mainly of calling attention to the fact that animals discriminate tilt, of which the writer has been well aware. A theory of shape recognition is unlikely to explain this capacity since a system picking out shape *per se* must discard information about tilt. If we suppose that there are two permanent lines on the two-dimensional array of cells (see Deutsch, 1955), one vertical and the other horizontal (each with its own final common cable) which send down a message when wave fronts from shapes externally imposed on the two-dimensional array arrive, then information about the tilt of lines will be available to the animal. This would solve the difficulty which Dodwell feels so much.

REFERENCES

- DEUTSCH, J. A. (1955). A theory of shape recognition. *Brit. J. Psychol.* **46**, 30-7.
 DODWELL, P. C. (1957). Shape recognition in rats. *Brit. J. Psychol.* **48**, 221-9.
 GELLERMANN, L. W. (1933). Form discrimination in chimpanzees and two-year-old children. I. Form (triangularity) *per se*. *J. Genet. Psychol.* **42**, 3-27.
 LASHLEY, K. S. (1938). The mechanism of vision. XV. Preliminary studies of the rat's capacity for detail vision. *J. Gen. Psychol.* **43**, 123-93.

(Manuscript received 2 February 1957)

PUBLICATIONS RECENTLY RECEIVED

The Measurement of Meaning. By C. E. OSGOOD, G. J. SUCI and P. H. TANNENBAUM.
Illinois: University Press. 1957. Pp. 342. \$7.50.

Take a sample of subjects and a sample of 'concepts', e.g. lady, boulder, sin, father: ask each subject to rate each 'concept' on a sample of 7-point bipolar scales, e.g. good-bad, robust-delicate, restless-quiet: sum the ratings of each 'concept', factor analyse, and rotate for simple structure, and you generally find the three orthogonal factors, namely, 'evaluative', 'potency', and 'activity' (some possible additional minor factors have not yet been fully worked out). Each 'concept' can be given a position in this 3-D (theoretically n -D) space, which represents its cultural meaning. Distance from the origin corresponds to intensity. This, in essence, is the method used by Osgood and his collaborators. As they point out, although every 'concept' has its own specifiable position in their semantic space, you cannot necessarily recover the concept from its three co-ordinates. To do so, you need additional information, the 'linguistic and situational context' in which the 'concept' appears. It is meaning *without context* which they claim to measure.

This seems to be a fairly obvious application of factor analysis, but one which is fraught with methodological difficulties, and which requires a good deal of spade work. It is to the credit of Osgood *et al.* that they have already accomplished so much. It is a pity that this is not more evident from a brief glance at the book; discussion and experiment are mixed together, and the reader cannot see at once which of the two any page contains. As we should expect with Osgood as senior author, the methodological difficulties are discussed frankly. For example, the authors note the failure of the seven steps on their rating scales to be of equal subjective size. They have estimated the amount of error to be expected from this failure, and have decided that it is sufficiently small to be neglected when compared with the noise level of their measuring instrument. Again, the authors note the failure of the bipolar dimension 'robust (rugged)—delicate' to pass through the zero of the evaluative factor (the whole dimension appears to be mildly 'good'). In this case they merely state that it is not possible to evaluate parts of a measuring instrument against other parts, without laying yourself open to the charge of circularity. They can as yet suggest no way of dealing with the problem.

Having frankly admitted these and other methodological difficulties, the authors proceed to ignore them, and describe some of the applications of their method. Success here is, in itself, a rough form of validation. They show that their method can give a more general description of attitudes than can be obtained by the use of the conventional single scale, which corresponds to their evaluative factor. They and their students have measured political attitudes, the conventional meanings of colours, and aesthetic judgements. They have also measured changes in attitude from exposure to information in various guises, the acquisition of meaning by novel 'concepts', and the temporary changes in meaning which occur when two incongruous concepts are associated. Roughly 60% of their more detailed predictions in this field have been borne out, and a number of novel findings have been recorded. Thus (p. 282) 'Whereas the meaning of NURSE can easily be devalued by attaching TREACHEROUS to it, attaching SINCERE to PROSTITUTE fails to budge the immoral connotation of the fallen woman.' This level of success is no mean achievement in view of our rudimentary knowledge of guiding principles in this area.

The authors and their students have also attempted to validate certain aspects of the Thematic Apperception and Rorschach Tests, and a number of hypotheses about personality and the content of dreams. Roughly 25% of the tests proved significantly successful. Finally they have made a detailed study of the three personalities in a case of triple personality, both towards the middle and end of treatment. With the minimum of background information, they made a clinical evaluation of the patient based upon their measurements. This evaluation does not appear to your reviewer to match the report of her two psychiatrists much better than by chance. However these rather unsatisfactory results need not necessarily be taken as a criticism of their method, for so many hypotheses in clinical psychology are incompatible that only a relatively small proportion are ever likely to be validated. And the reports of psychiatrists are normally loaded with subjective errors. In fact, only a relatively small amount of agreement in this clinical area is to be expected, if their approach is a valid one.

This book describes the first really successful approach to the experimental study of meaning. What conclusion do we reach? Meaning is simply three main dimensions (in a hypothetical 'semantic space')

combined with a context. At first, this may seem a little difficult to accept, given the amount of attention which meaning has received from philosophers and from some of the older psychologists. Are our doubts any different from the doubts which are always expressed by contemporaries when a longstanding mystery is explained away? I wonder.

E. C. POULTON

The Psychology of Perception. By D. W. HAMLYN. London: Routledge and Kegan Paul. 1957. Pp. vii + 120. 12s. 6d.

In this book Hamlyn attempts a philosophical critique of the views of the Gestaltists, Michotte, Gibson, Hayek, and the Transactionalists on perception. His main thesis is that these writers have not been offering, as they thought, scientific explanations of perception, but have merely offered different conceptual analyses of what perception is, and that these analyses are in all cases mistaken. He maintains that a correct philosophical analysis of the concept is an essential prerequisite to any scientific explanation of perception. Unfortunately he makes no attempt to supply a correct analysis of perception himself, and the result is that the tone of the book is completely negative. It is by no means obvious that a correct philosophical analysis of a concept in ordinary usage (or rather—in the case of ‘perception’—in philosophical usage) is a prerequisite to a scientific explanation of the phenomena covered by the term—Hamlyn is content with a dogmatic assertion, he does not argue the point: if the development of chemistry had had to wait upon correct philosophical analyses of words such as gold or heat, we might still be in ignorance of atomic theory to-day. What is important to the scientist is to be clear about the facts which he is trying to explain. Koffka’s discussion of the question ‘why do things look as they do?’ is an attempt to get clear about these facts, and the useful part of his discussion is that an answer to the question must start from the physical stimulus as it affects the retina and not from the object itself.

Many of the arguments Hamlyn deploys against psychologists are put forward with such dogmatism and cursoriness that it is difficult even to make a guess at what could have led the author to present them. Thus, on page 19, he maintains that it is impossible to explain perception in terms of underlying brain processes because perception involves learning and learning cannot be explained in terms of a brain process: no reason is given for this curious belief. His main thesis against the Gestaltists seems to be that it is impossible to give explanations of normal perceptions (p. 31)—this is simply the way we perceive when there are no abnormal conditions which make us perceive abnormally. Yet any full explanation must explain both the normal and the abnormal cases. A match normally lights when struck: a full explanation will explain both why it normally lights, and why it fails to light if conditions are abnormal (e.g. if there is no oxygen in the air). What the psychologist is trying to do is to formulate the relationship between stimuli and responses and to make hypotheses about the intervening mechanisms: yet Hamlyn has succeeded in writing a book on this topic with scarcely a reference to responses and with no discussion of the term.

Hamlyn is consistently condescending and unfair to the psychologists he discusses. He makes no attempt to see what is useful about their views: e.g. it is quite true that Hayek uses the word ‘classify’ loosely both of the whole organism and of the nervous system. Hamlyn rightly points this out, but he does not go on to consider what Hayek was really trying to say and whether it is a useful way of looking at things.

The topics discussed in the book seem to be chosen almost at random: there is no mention of Brunswik, no discussion of explanations involving unconscious inference. Nor do some of Hamlyn’s own mistakes help to inspire confidence: it is alarming to find that the author of a book entitled *The Psychology of Perception* believes that light from objects low in the visual field falls on the bottom of the retina (p. 88). Moreover, Hamlyn’s own philosophical confusion seems to be at least as great as that of the psychologists he decries.

Since psychologists often do fall into conceptual confusions in discussing perception, a book written by a philosopher on the psychology of perception could have been extremely helpful to psychologists. To be helpful, however, the author would have had to make a genuine attempt to see what psychologists have been trying to say, rather than be content with trying to expose verbal confusions in their attempts to say it.

N. S. SUTHERLAND

The Scientific Study of Social Behaviour. By MICHAEL ARGYLE. London: Methuen. 1957. Pp. viii+239. 21s.

Mr Argyle intends this book for students of Psychology and of the other Social Sciences. He distinguishes it from a text-book on the grounds that it deals selectively with the field and, for instance, omits discussion of socialization and of personality studies. In this book, the author wishes to put facts before theories and then to examine the various theories in the light of these facts. This procedure, Mr Argyle considers, differs from that of text-books of Social Psychology, which he describes as 'often illustrating a particular theoretic viewpoint with a congeries of experiments, quasi-experiments, and other people's opinions'. One wonders whether this description fits some of the better text-books in the field, e.g. Asch's *Social Psychology*. Mr Argyle further hopes that the book will be of use to social administrators by indicating methods of research which would help in the solution of practical problems.

It is on these grounds, then, that the book should be evaluated. It begins in a very promising way with a lively introduction in which Mr Argyle stresses that the study of social behaviour is 'a field of experimental psychology with a special subject matter'. Here, however, he limits it unduly to the study of the behaviour of people when others are present. An equally important sphere of study for the social psychologist is that of the influence on the individual of those who are not present, for example, the influences of mass media and of wider group belongingness.

The book is uneven. Some parts, like those concerned with small groups, are excellent. A wide range of research material is presented in a vivid manner and in sufficient detail to permit the reader to follow Mr Argyle's line of argument. This is not always the case in those sections in which Mr Argyle is perhaps less interested. In these sections, researches are listed but the data are presented in such a concentrated manner that it is difficult to make an assessment; also, the style tends to be more abrupt and, particularly in the first section, dealing with methods of studying social behaviour, the author does not always make it clear where description of the data ends and interpretation begins. The section dealing with the design of social investigations is well organized; in it, the author discusses a wide range of empirical designs and examines each in turn to see how far they satisfy the conditions of valid scientific inference. Chapter III, 'The Explanation of Social Behaviour', lists the various stages in theory construction. This section is very interesting. When Mr Argyle passes from the consideration of theories to generalizations, the book becomes a little too much of a catalogue. Concept after concept is described without always informing the reader of the link between them. For example, in a short section entitled 'Self Perception', we are told about accuracy of perception, about insight, about level of aspiration, about self perception, about congruence between different aspects of self perception after therapy, and then are taken back once again to level of aspiration.

The over-all impression gained is that this is an important reference book for the student of Psychology, first for literature references over an unusually wide field of social psychological problems, and second for the novel and stimulating approach of some of its sections. In particular, one would recommend those sections dealing with the design of experiments. The book has not, however, quite achieved what it set out to do. It has not integrated facts—indeed it suffers from the lack of Gestalt that theory helps to impose on data. Part of the difficulty arises from the sheer vastness of the field, part from the varying degree of the author's 'at homeness' with the material. I look forward to an intensification of Mr Argyle's obvious interest and knowledge in the study of small social groups, which should result in another book. Meanwhile, I should like to recommend his book, with the reservations mentioned, to all those who wish to familiarize themselves with recent research in the field of Social Psychology.

HILDE T. HIMMELWEIT

Memory: Facts and Fallacies. By IAN M. L. HUNTER. Harmondsworth: Penguin. 1957. Pp. 185. 3s. 6d.

As most psychologists are aware there is no single volume which covers the various scientific studies of memory. Gomulicki's British Psychological Monograph represents a laudable attempt to classify the various theories which have been put forward and it is to be regretted that so far no author has attempted to do the same for the experimental data.

Dr Hunter's book, published in the Pelican Psychology Series, is naturally designed for a wide public but in many respects it is an attempt to present the experimental data; its title might well be 'Memory: Facts, not Fallacies' for Hunter is not out to expose but to present the evidence. This he has done in as straightforward, matter-of-fact an account as could be demanded. The plan of the book is to make some

appraisal of what we understand by memory, to describe the processes of memorizing and discuss the nature of forgetting and recall. This is followed by chapters on the repressing and imaging of material, and a final section is devoted to 'improving memory' where the discussion is predominantly about mnemonic systems and their relation to other efforts to classify material.

Within this framework there is no doubt that Hunter has marshalled many of the accepted experimental data of remembering. We have here a workmanlike little text and one that has studiously avoided the novel and sensational. Much recent work, particularly in the fields of short-term memory and neurology, seems to have been deliberately excluded. The most controversial chapter is that on 'Repressing' where, after indicating the influence of repression on free association, an effort is made to explain the loss of childhood experiences along conventional lines without resource to psychoanalytic interpretations. Even here the author is quick to acknowledge that we have 'too few facts and too many question marks' to decide between the processes.

Some psychologists will I think find the pedagogical tone of the book a little irksome. It is admirably direct but it can be dictatorial. Thus we are told at the beginning: 'The basic memory processes are four in number, namely, learning, remembering, forgetting, and retaining' (p. 14). This is an odd quartet, whilst those readers who think that the initial perceiving is an integral part of the process will find its omission puzzling, particularly as Hunter is keen to indicate Bartlett's contribution.

But just because this text is so unassuming and direct it does raise a major query for the psychologist. Hunter himself believes: 'The great advances of psychology in the present century have been due to the adoption of the scientific method' (p. 136). How far we consider this to be the whole truth will probably depend on what we understand as scientific method, but since this is an account for the general public of what has been accomplished in this field we, as psychologists, may well ask ourselves: 'How much has been discovered?' 'What is the nature of any advances which have been made?' 'Where do we go from here?' The point is made all the sharper by Hunter concluding his book with Fulwood's 1562 poem on 'Memorie' which prescribes amusingly close to the present-day scientific formula.

It has to be admitted that the advances are not spectacular. In the reviewer's opinion, psychologists have observed two major processes which would help us to explain the aberrations of remembering. First, our present remembering is inaccurate because in the past, due to our own individual experiences, we did not perceive and assimilate accurately (Bartlett), and secondly, something which is close to the other side of the coin, our remembering is inaccurate because a present repression is inhibiting a past percept (Freud). Evidence would suggest that Bartlett's is much the more ubiquitous experience but there seems no reason for denying the occurrence or import of the Freudian phenomena. Further advances are slight: there are several variations on the Bartlettian theme which have supported and slightly extended his original monograph, whilst the work with brain-injured amnesia patients has brought out the curious time-relations which exist between experiences. Recent studies on short-term memories have also emphasized both this and the differences between immediate and long-term recollections. But *in toto* it is a slim contribution; on the one hand the neurological picture is inevitably confused because of the difficulty of ascribing with such a complex and interacting system any one phenomenon to any particular neurological defect, whilst any general theory of remembering is inadequate because it inevitably becomes involved in the whole gamut of mental processes about which insufficient is known. I think we have to accept that future work on remembering must tackle this problem in a piecemeal and systematic fashion and it is of note that American studies in this field do attempt to do this much more than we do in Britain. It would seem to be a reminder that when we are confronted by such processes as remembering there is a need for more systematic and sustained attacks upon the problem if we are to progress.

HARRY KAY

Personality, Appearance and Speech. By T. H. PEAR. London: George Allen and Unwin. 1957. Pp. 167. 21s.

Prof. Pear defines personality as 'the effect upon others of a living being's appearance, behaviour, etc. so far as they are interpreted as distinctive signs of that being'. A tremendous increase in contacts, direct or indirect, between people has taken place in our age of 'mass communication'. The social importance of judgements of others (based on speech, dress, mannerisms, etc.) is greater than it has ever been before; partly because of the 'effects upon others' that an individual may exercise when suitably processed by those who have recently been called the 'hidden persuaders'. What do psychologists know about all this? Very little, says Prof. Pear, and he proceeds to fill the gaps, chapter by chapter. Dress, speech, face, gesture, gait—hundreds of subtle and amusing observations follow each other. At the end of it all, the

reader is left breathless and frustrated. The book is curiously disjointed: it is as if we were invited to listen to a brilliant after-dinner conversation that Prof. Pear is having with himself. The titles of chapters give an indication of the general theme for the next few pages; what follows is a series of variations on the theme, each of them enjoyable, but only a few consecutively arranged and developed in the grim pursuit of statements of general validity.

Prof. Pear is a keen observer of the social scene. Many of the almost self-contained paragraphs are full of observations which are a challenge to further research, and will have to be taken up, sooner or later, by those interested in social perception. Neglected sources of information are often brought to light. For example, why is it that psychologists have been so little concerned with the significance of differences between national groups in their way of describing similar situations? 'The English "cope with" a difficulty, implying that success may not be either immediate or complete, Americans, more optimistically, "fix" it.' An Englishman 'stands for Parliament'; an American 'runs for Congress'; An Australian says 'Too right it is' when an Englishman says 'Not half it isn't'. Or, a glimpse into a musty corner: 'Teachers in "finishing" schools and deportment classes will have lists of "do's" and "don'ts"; it would be interesting to learn how these change with time; who initiates improvements; whether they are ever resisted.' Similar indiscreet and pertinent questions are asked about beauty contests, male and female, about the ups and downs of radio, television and film 'personalities', and about innumerable other fads which are suddenly with us, as by some miracle—often carefully contrived.

Concluding his book, Prof. Pear writes: 'Omelettes cannot be made without breaking eggs: in this book several new ones have been cracked: the noises may have surprised some psychologists more than the ordinary reader.' The omelette was certainly well prepared; but this reviewer begs to disagree with the author's estimate of the number of eggs; it was not a substantial meal, and the noises were not very surprising. Prof. Pear's concern is with a problem, perennial in psychology, of the relationship between controlled investigations, conducted in the usual ways, and the rich complexity of 'real life'; he finds that there is a large discrepancy between the two. Very often this is true; the most promising way out is probably through isolating variables of higher order of complexity and applying to them, as far as possible, methods which have proved useful in simpler contexts. This is being attempted all the time, with varying degrees of success. This book will certainly stimulate some psychologists to turn their attention to some new and complex variables and, in doing this, it will perform a valuable service. But it does not go very far beyond demonstrating that complex variables do exist, and it does not contain many hints as to what should be done about it.

H. TAJFEL

A Short History of Psychotherapy. By NIGEL WALKER. London: Routledge and Kegan Paul. 1957. Pp. xii + 185. 25s.

This book is intended mainly for non-psychotherapists with an interest in the subject. It traces the development of modern methods from Freud and Jung right down to the present day, including such topics as group psychotherapy and therapeutic communities. It is short and blissfully lacking in that semantic vagueness and exuberance which characterizes so much writing in this field. The author is not afraid to tackle such topics as the efficacy of psychotherapy, a point taken for granted by most writers. Altogether, this is an excellent book very much to be recommended to those interested in the topic. On the critical side, I found somewhat disturbing his omission of several people who have made very important contributions—more so on any reasonable appraisal than others who are mentioned. The index of names refers to Hiawatha, but not to Herzberg whose doctrine of 'active psychotherapy' was an important development. Wolberg is mentioned, but no Wolpe whose use of learning theory in connexion with psychotherapy has become one of the most important new developments in this field. However, with these exceptions, the book is remarkably complete considering the small number of pages. Presumably the author might argue that opinions and the importance of individuals differ and will continue to do so.

H. J. EYSENCK

The Sane Society. By ERICH FROMM. London: Routledge and Kegan Paul. 1956. Pp. xiv + 370. 25s.

This book is not so much a psychological study as a tract for the times describing, somewhat after the manner of the sage of Ecclefechan, first, the parlous state into which human society has fallen, and secondly, the conditions under which humanity may escape perdition if only it will heed the prophet's

warning. The author's approach is, of course, that of the American school of culture psychologists, who, in an attempt to escape the limitations of Freud's biological approach, have directed attention so predominantly to the influence of cultural factors on individual development as to surrender the insight into unconscious motivation achieved through Freud's researches. It is a guiding principle of the author that the problem of psychological abnormality is not so much one of failure on the part of the individual to adjust to the conditions prevailing in a given society as of failure on the part of society to provide an answer to man's existential problems; and it is in the light of this principle that he feels justified in applying the terms 'sane' and 'insane' to social systems. Man's existential problem is further regarded as arising out of his having lost that harmony with 'nature' which he originally enjoyed through instinctive adaptation, and his experiencing a consequent necessity to make good the loss of this original harmony by seeking 'ever-higher forms of unity with nature, his fellow men and himself'—a necessity which is 'the source of all psychic forces which motivate men'. In the author's view, the freedom conferred upon mankind by the dissolution of his pristine harmony with 'nature' becomes a source of considerable anxiety to the individual who is, accordingly, only too liable, in his attempt to achieve security through some new harmony, to allow himself to become the puppet of the prevailing social system—whether this be (in terms of the contemporary situation) a capitalist Western society in which the individual is treated as a commodity or a Communist society in which he is treated as a slave. The only escape from the 'robotism' and consequent 'insanity' involved in both these prevailing forms of society is, in the author's view, that provided by 'Humanistic Communitarian Socialism'—which would appear to be a social system in which the individual is treated as an end rather than a means, is afforded an opportunity for creative activity, and is related to other individuals by bonds of love. The general conception of such an ideal society is, of course, not entirely novel, having been already described in the New Testament under the less cumbrous title of 'The kingdom of heaven'.

Whatever interest Fromm's theme may possess at a superficial level, his treatment of it displays the characteristic weakness of the American school of culture psychologists in that, adopting the approach of 'humanistic psycho-analysis', he seeks to describe human behaviour in terms of object-relationships while at the same time ignoring the concepts of internal reality and relationships with internal objects. His ignoring of these concepts renders it easier for him to entertain the notion of man's pristine harmony with, and subsequent loss of, 'his original home, nature' (better known perhaps as 'the garden of Eden'), while apparently remaining oblivious to the unconscious significance and motivation of the phenomena to which he refers.

W. RONALD D. FAIRBAIRN

Psychoanalysis and Psychotherapy. By FRANZ ALEXANDER. London: Allen and Unwin. 1957. Pp. xiv + 299. 25s.

Dr Alexander, a native and medical graduate of Budapest, became a leading member of the first training institute for psychoanalysts soon after its foundation in Berlin in 1920. Since he emigrated to the U.S.A. in 1930 he has been an influential member of the American Psychoanalytical Association, particularly through his work at the Chicago Institute of Psychoanalysis and as Clinical Professor of Psychiatry at the University of Illinois. His comments upon contemporary psychoanalytic theory and practice in the U.S.A. are therefore both authoritative and interesting.

Of the essays which have been welded together into this book, the first group is concerned, broadly speaking, with the problem of explaining what takes place in the therapeutic process of psychoanalysis. The old fallacy that it consists of the extension of the patient's self-knowledge is discussed once more (perhaps this is necessary in a country where Adler's influence was so strong); and the importance of corrective emotional experience and of the patient's relationship to the analyst is stressed in a way which to British psychoanalysts will seem entirely orthodox. Dr Alexander has a distinction to make between two kinds of regression, one of which seeks an earlier stage of development because it was so satisfactory, while the other does so because the later stage was so unsatisfactory; but he does not draw any very practical conclusions from the distinction, which seems to me merely to separate two sides of the same penny. Perhaps his most practical observation is that the orthodox Freudian ideal of a wholly impersonal analyst, whose own reactions to the patient are kept to a carefully controlled minimum, is not only unattainable but also a positive hindrance to therapy; this is one of the unacknowledged ways in which post-war Freudians on both sides of the Atlantic have moved a little toward the Jungian point of view. (Its danger is, of course, that the moral judgements which orthodox Freudianism succeeded in banishing will be smuggled in again under a professional pseudonym.)

The second group of essays is concerned partly with the relationships between the pure salt of psychoanalysis and the various solutions of it which Dr Alexander labels 'psychotherapy', and partly with the problems of training practitioners of both kinds. Some of these problems are less real in this country than in the U.S.A., where psychoanalysis gained a quicker and wider acceptance in all fields of psychiatry between the wars (it has been suggested that the chief reason was the gap left by the slower development of physiological techniques of treatment). However this may be, it almost sounds as if one of the hardest tasks of psychoanalysts in the U.S.A. from now on will be, not to secure popularity for their ideas, but to maintain a pure strain from which to revitalize the many cross-bred techniques that are reproducing themselves.

NIGEL WALKER

Introductory Clinical Psychology. By S. L. GARFIELD. New York: The Macmillan Company. 1957. Pp. xiii + 469. 42s.

The purpose of this book is to give beginners a picture of test instruments in the field of intellectual functioning and personality. In addition, it discusses the role of the clinical psychologist as research worker and therapist. There is a very general first chapter which gives what is meant to be a history of clinical psychology and a final section dealing with scientific and professional problems. Discussions of such matters as validity and reliability are not misleading because they are so general, but they do not enter into any of the complexities of the problems. This reviewer considers that such complexities might be hinted at even in an elementary text. In its discussion of tests, the book seems adequate in so far as it covers most of the recognized and widely used clinical material from America. But it shows little awareness of other material apart from the Rorschach: thus, Eysenck's work with personality tests is not mentioned at all and Shapiro's study of brain injury is ignored. This may account for the undue weight given to projection tests.

Nearly 100 pages of this more than 400-page book is given to treatment (as usual called therapy). The author then notes that there are comparatively few studies devoted to evaluating the outcome of therapy. The word 'comparatively' is ambiguous but, since the First World War, the number of such studies is by no means inconsiderable especially in view of the uniformity of their findings. It is wrong of Prof. Garfield to ignore the consistent failure of these studies to prove the value of psychological and psychiatric treatment. It is a pity that the systematic ignoring of such data on the part of clinical psychologists was not emphasized in this text. The reason, however, may be that the author proceeds to ignore the findings of the one review he mentions.

The last few chapters of the book are concerned with the American psychologist and his professional problems and are unlikely to be of much interest to clinical students here. It is hard to praise this book, but there is a relatively lively and critical chapter on research which shows that the unfortunate state of clinical research is well understood by the author just as his book exemplifies the unfortunate state of clinical practice.

N. O'CONNOR

Physiology of the Nervous System. By E. G. WALSH. London: Longmans, Green. 1957. Pp. xvi + 563. 50s.

There has long been a need for a neurophysiology text which introduces and organizes the facts about the *inside* of the organism for those who are concerned with the behaviour of that organism. This book is welcome not only because it is aimed at psychologists, among others, but because it incorporates recent developments, both of fact and fancy, within the field of neurophysiology. In addition to the traditional headings of 'Nerve and Muscle', 'The Visual System', 'The Cerebellum', the reader will find chapters on 'The Applied Physiology of Pain', 'The Sensory Pathway, Thalamus and Parietal Lobe, Speech', with frequent reference to material with which the psychologist will be comfortably familiar—such as the tilted room experiments of Witkin, the inverted field experiments of Stratton, the descriptions of postural development by Gesell. The book is profusely illustrated, sometimes to the extent of containing figures for which no reference can be found in the text. The psychologist is introduced to topics of current excitement and controversy in neurophysiology—the brain stem reticular formation (although this receives less attention than it would appear to merit), the 'visceral brain', the efferent control of sensitivity of muscle spindle receptors, the work of Weddell and co-workers on somatic perception. The author presents relevant applications of modern servo theory to some of the subject matter. Frequent and intelligent reference is made to neurological case material.

It is not by way of adverse criticism to note that the reader will not find in this book a *systematic* treatment of neurophysiology. It is more a collection of interesting facts and speculations emitted, so it would appear, as they occurred to the author. Thus, we shift abruptly from the structure of the insect eye to the unidirectionality of synaptic conduction. Even in the chapter headings there is no attempt to impose a rigid organization.

Since the book is 'intended to fill the gap between the physiologist on the one hand and the clinician and psychologist on the other', questions arise which apply to all gap-fillers. Can a gap be filled when the boundaries on either side are not very firm? Assuming firm boundaries, how safe is the bridge connecting them? On both questions the book leaves something to be desired. The author's attempt to be 'psychological' is laudable, but often his psychology is not as firmly established as even that of the introductory student of the subject, who would not confuse 'instrumental' or 'operant' learning with Pavlovian conditioning, nor depend frequently on concepts such as 'sensorium', nor make unqualified assertions such as: 'The human being uses many visual cues for orientation, the rabbit none' (p. 156). Or again 'In the human being it seems that what can and cannot be done [posturally] has to be discovered by *experiencing in infancy*' (pp. 161-2, italics in original). Or 'Conditioning is most readily achieved if there is a simple relationship between stimulus and reinforcement. These principles are perhaps the most enduring feature of the work on conditioned reflexes' (p. 221). It is fruitless to argue who has the greater responsibility to learn the other chap's subject—the psychologist or the neurophysiologist. Nevertheless, it is the case that psychology can (and in the opinion of many, such as Skinner, should) exist without reference to neurophysiology, while neurophysiology (and more especially neurology), in so far as it attempts to relate the behaviour of parts of the inside of the organism to the behaviour of the whole organism, cannot escape at least a description of such behaviour. Psychology has, over the past twenty years, developed both descriptive systems and behavioural theories which could prove useful to those in related disciplines who wish to talk of behaviour.

It is also clear that there is a large gap between the gap-fillers in neurophysiology and the gap-fillers in psychology—even when they are studying precisely the same phenomena. The present book—or any other neurophysiology text—really does not close that gap. Thus, while the author has discovered Lashley, he does not—nor do most physiologists—read the journals in which psychologists customarily publish their experimental contributions to neurophysiology (notably the *Journal of Comparative and Physiological Psychology*). Thus, in discussing the frontal lobe syndrome there is no mention of several important papers by Harlow, Settlage, Pribram, Mishkin, and others. Equally important work on the effect of temporal neocortical lesions and limbic system lesions receives no mention. The extensive work of Beach on the cortical control of sexual behaviour is not noted. This list could be extended and elaborated indefinitely. In each case a reference to material in psychological journals would have contributed *directly* to topics already discussed in the text and, more important, would sometimes have altered the argument significantly. At the same time this book makes a much greater effort to bridge the gap than do most neurophysiology texts, and it is safe to predict that it will be referred to frequently by students and specialists alike.

LAWRENCE WEISKRANTZ

Extrasensory Perception. By G. E. W. WOLSTENHOLME and E. C. P. MILLER (eds.). London: Churchill. 1956. Pp. ix+240. 27s. 6d.

The papers and discussions which make up this notable Ciba foundation symposium of 1955 offer material to suit many tastes. Most of the twenty-three symposiasts live in Britain, but visitors from France, Holland, Spain and the U.S.A. also took part. Many hold scientific posts, some are medically trained and at least one is a philosopher. Partly because so much of the book is taken up with the discussions, it will be more profitable to give the impressions made by the book than to present something like a table of contents. The foremost impression is of the variety of attitudes to extrasensory perception which can be adopted by men of parts, all willing to listen to what the other man has to say. One finds, for instance, the attitude that the physicists must now rewrite their systems to accommodate the facts of para-psychology, a rewriting analogous perhaps to that forced by the revelation of the difficulties inherent in the concept of simultaneity. Then there is the view that physics need not be rewritten, but that physiology must admit some hitherto unrecognized sensory mode. Others again require no rewriting of theories; it is by training, deliberate or otherwise, that skill in particular directions is attained by individuals, equal to that of the best conjurers, who can, however, be relied on to simulate or reproduce practically any phenomena yet observed. Another view is that there is nothing to explain and that the

extraordinariness of the phenomena is (like beauty) in the eye of the beholder. A fifth is that science need not be rewritten to accommodate the phenomena, since they are different in kind from physical events and do not proceed within the physical domain; and finally some participants would, I think, adopt the position that odd and interesting things happen, which we do not yet understand but which will in due course be resolved by the application of scientific method.

When all is said and done, I can discover no hint that any member of the conference modified his attitude in any degree as a result of what was said. No comment on Dr Wasserman's 'fields' suggests to me that his hearers had understood him. Pigeons may or may not be clairvoyant, but Dr Matthews seems at the end to be no less in the dark as to how clairvoyance could be of service to get them home; and, for all Dr Parkes's optimism, many a glass of sherry will need to be downed before the difference between Dr Amadou and Mr Spencer Brown is dissipated.

Finally, there remains to puzzle us the contrast between the complexity and meaninglessness of the controlled situations in card guessing experiments, on the one hand, which leave the layman baffled, sceptical and unconvinced and, on the other, the spontaneous phenomena, so pregnant, so unforeseen, and so seldom repeated as to arouse in the scientist the worst suspicions. The former type of phenomenon has become a battleground for intellects; it is from the latter that the study of extrasensory perception renews its inspiration and draws its sustenance.

B. BABINGTON SMITH

Some Recent Developments of Psychology in Great Britain. By F. C. BARTLETT. Istanbul University. 1957. Pp. viii+92. 2.65 T.L.

In 1956 Sir Frederic Bartlett gave a series of six lectures at the Faculty of Letters, Istanbul University, which documented some of the major developments in British (one might almost say in Cambridge) psychology over the past forty years and highlighted the social changes which brought them about. Two further lectures were delivered at the Technical University of Istanbul and repeated in Ankara to largely military audiences. All eight of these lucidly brief lectures are printed here, first in Turkish and then in English. The lectures are admirably suited to their purpose of giving a clear picture of what an important cross-section of British psychologists have been doing in recent years, of how they have set about doing it, and of why. Any historian of the present century's psychology would be unwise not to read this little book.

I. M. L. HUNTER

Ciba Foundation Colloquia on Ageing. Vol. 3: *Methodology of the Study of Ageing.* Ed. by G. E. W. WOLSTENHOLME and C. M. O'CONNOR. London: Churchill. 1957. Pp. x+202. 32s. 6d.

In 1954 the Trustees of the Ciba Foundation decided upon special measures for the encouragement of laboratory and clinical investigations relative to the problems of ageing. Three conferences—or 'colloquia'—have been organized with this particular aim in view. This volume contains the proceedings of the third colloquium in this series. The contents range from a paper on 'The Study of the Ageing of Cells', by J. F. Danielli, to 'Twin Data on the Genetics of Ageing', by Franz J. Kallmann. Three of the twelve papers are likely to be of special interest to psychologists.

F. Verzá, of the Physiological Institute at the University of Basle, offered an interesting emphasis on 'adaptation' as the most suitable basis for gerontological research. This seems absolutely sound, provided one sticks closely to the definition of adaptation as 'the general capacity of living organisms to live under continuously changing conditions', and this reviewer is in close accord with Verzá when he states: 'When this adaptability decreases, ageing begins.'

Irving Lorge, of Columbia University, New York, presents a competent survey of the methodological problems faced by the psychologist in the study of intelligence and emotion in ageing. One is startled, however, to read that 'intelligent behaviour may be conceived as the capacity to make good, adequate and acceptable responses'.

The third paper is by A. T. Welford, of Cambridge, and is a model of its kind. He deals with methodology in the study of changes in human performance, and in so doing draws attention to point after point the importance of which it is difficult to exaggerate. It is to be hoped that every psychologist newcomer to the investigation of changes with age will have this paper drawn firmly to his attention.

Although the book includes the full discussion on each paper—sixteen people were present in addition to those giving invited papers—this but serves to remind us that the indefatigable organizers of these

top-level occasions have a very hard task indeed. This said, let us express the sincere hope that they will go on trying: the potentialities of the small colloquium appear far greater than do those of some large congresses.

ALASTAIR HERON

Human Relations in Interracial Housing. By D. M. WILNER, R. P. WALKLEY and S. W. COOK. Minnesota University Press. London: Cumberlege. 1956. Pp. xiv + 167. 32s.

This book is an attempt to assess the relative influence of living in integrated or segregated housing projects on the attitudes and behaviour of white Americans. The research was carried out on two pairs of housing projects which were as alike as possible except that one in each pair had an integrated, and the other a segregated pattern of living. Within each area the relations between proximity, and the variables of contact, social climate and ethnic attitudes were studied. As there were only 10% of negro residents in each project the concepts of 'far' and 'near' were applicable within each separate housing estate and, thus, an added check was obtained. In general, it was found that contact with negroes did influence the attitudes of the white people towards them. 'Whether we consider the initially more favorable or initially less favorable respondents, those who live near Negroes in a project are more likely than those living farther away to report neighborly contact, to anticipate that white friends in the project will approve of such contact, to have high esteem for the Negroes in the project, to approve of the biracial aspect of the project, and to have a favorable attitude to Negroes in general. These "near-far" differences may be interpreted as reflecting change in attitude at least for "nears". It is of interest to note that such changes take place even among those with an initially less favorable attitude.'

This book is a valuable contribution to the growing volume of work in social psychology done on people as they are in a real-life context. In addition, the writers have tried to apply the rigorous methods of experimental analysis. It should be of interest both to people working on race relations and group dynamics, as well as to people engaged in administering mixed racial groups.

MADELINE KERR

Modern Sociological Theory. Ed. by H. BECKER and A. BOSKOFF. New York: The Dryden Press. 1957. Pp. xiv + 756. \$6.50.

The index of this huge book shows that Max Weber is referred to on 107 of the 760 pages of text, together with other early European sociologists like Durkheim, von Wiese and Gurvitch. American sociologists who were born or trained on the Continent also receive frequent mention—Becker (161 pages), Sorokin and Znaniecki, together with Talcott Parsons who has translated and interpreted the early writers. A glance at the list of twenty-six contributors shows that none of them is at what are generally considered to be the main centres of sociology in America—Chicago, Columbia, Michigan and Harvard. All this suggests that the book is not representative of the main stream of current American sociology, but of that minority group whose members look back to pre-1914 European thought. They would not approve of the above statistics, since they prefer more informal data which their theories help us to understand, not to predict. They might approve the following description of what we may call the 'Theorist' type of sociologist, which we shall contrast with the 'Predictor' type of sociologist with whom we are more familiar. (a) Theorists try to gain some understanding of phenomena by dividing them up in illuminating ways into different types of groups, leaders, etc. (b) They are critical of 'empiricism' on the grounds that it produces mere collections of facts, leads to unimportant research, and diverts attention from the important theoretical issues (pp. 190ff.); they prefer suggestive, though less precise information, such as can be obtained from case studies and historical research. (c) They are interested in the sociology of knowledge, law, religion, art and (for some reason) peasants. Predictors have a taste for cities, industry, small groups and public opinion. Both are concerned with social class and the family. (d) Theorists are preoccupied with logical and metaphysical matters on which empirical data have no bearing; they are concerned with the historical origins of their ideas and they write in a scholarly and verbose manner.

The development of such typologies, or classifications, is important, and it does add to our understanding. However, the theorist takes the problem no further—the enlightening classification is his final goal. He does not go on to inquire, for example, what the conditions are which make a sociologist a Theorist rather than a Predictor. For the Predictor, such classifications are an essential first stage, he then proceeds to find the empirical laws describing functional relations between variables so described.

However, there are certain empirical questions involved in constructing typologies which theorists do not face squarely: How high, for instance, are the correlations between the various elements in the typology? Not all the contributors to the present book are opposed to orthodox empirical methods—to continue our analogy. Predictors generally proceed from typologies to dimensions, and have empirical methods for deciding which set of dimensions is the best.

What is sociological theory anyway? The theorists regard their classifications as theories. Predictors are mainly interested in theories as a source of predictions to be tested—and no predictions follow from the Theorists' theories. There is little in sociology to correspond to the theories in psychology or the natural sciences—partly because there are so few laws to explain, partly because no one is agreed on what sociological theories would be like, and because neither Theorists nor Predictors are interested in deductive theories. One possibility is explanation by reduction to psychological findings, and a few miniature systems of this sort have been constructed, mainly in the small group area.

The first part of the book, 'Major Strands in Theory and Methodology', expounds the Theorist position for 300 pages, with much historical reference: the present reviewer did not find it very illuminating. The following section 'Some Specialisations in Modern Sociology' has a good chapter on religion by Honigsheim, together with the expected ones on knowledge, law and, surprisingly, small groups. The dead hand of the Theorist falls more heavily on some of these chapters than others. 'Convergences of Bordering Fields with Sociology' contains an interesting chapter on psychoanalysis by Hinkle. The final section 'Sociological Research and Theory Abroad' has an interesting account of British sociology by Professor Sprott, and also covers France, Germany, Italy and Japan; this is the most interesting part of the book.

The book is beautifully produced, contains a wealth of material on a variety of subjects, has extensive bibliographies and a good index, and is reasonably priced.

MICHAEL ARGYLE

How Children Learn: an Educational Psychology. By A. N. FRANDSEN. New York and London: McGraw-Hill. 1957. Pp. xiv+546. 41s. 6d.

This is a workmanlike text by an author who is familiar with both the academic literature and the school-room and who never loses sight of his subject-matter, namely, the real-life child in process of learning. In contrast to several recent texts, it avoids excess of, on the one hand, arid ivory-tower theory and, on the other, banal superficial anecdote. The bulk of the book is devoted to the relations between learning activity and the seven factors of maturity, teacher-guidance, practice, perception of the effects of provisional tries, transfer, motivation, and emotional disturbances. Then follow discussions on appraising individual differences in abilities, providing for these differences, and appraising school achievement. For the British student undergoing teacher training, the text's usefulness may be limited by its American setting. It should, however, be of great assistance in providing orientation and content to those concerned with giving courses in either educational psychology or the psychology of learning.

I. M. L. HUNTER

Sense and Nonsense in Psychology. By H. J. EYSENCK. London: Penguin Books. 1957. Pp. 349. 3s. 6d.

This is good light reading and will appeal widely by reason of its style, which is lucid and intelligible, and its subject-matter, which is frankly popular—hypnosis, lie detectors, telepathy, dreams, personality, politics, aesthetics. Prof. Eysenck is always interesting, often entertaining, occasionally irritating. When he gets on to psychoanalysis his critical level sadly deteriorates—any kind of tendentious argument or trivial anecdote will do to belabour his *bête noire*. This is a pity: every sensible psychologist is heartily sick of this tiresome guerrilla warfare; and there is no need for it, since psychoanalysts lay themselves open in plenty to serious and responsible criticism. Fundamentally I believe Eysenck is much too confident about things. There is not perhaps quite such a sharp and certain division between sense and nonsense as he would make out. It is somewhat surprising to be told that the main outline of the picture of human personality supplied by Pavlovian theory is unlikely to require any revision! Are the foundations of this theory so very certain? In the history of science the sense of yesterday not infrequently becomes tomorrow's nonsense; and nonsense turns out after all to be sense. I suspect the future development of psychology holds some surprises in store; and perhaps some major revisions of outlook after all! But breezy confidence makes good popular fare!

L. S. HEARNshaw

Alfred Adler: Apostle of Freedom. By PHYLLIS BOTTOME. (3rd ed.). London: Faber and Faber. 1957. Pp. 300. 25s.

Adler's stock is rising, so the third edition of Phyllis Bottome's biography, first published in 1939, has appeared at an opportune time. The author, a novelist, was an ardent disciple and close personal friend of Adler, and she gives a vivid account of his life and personality. To psychologists and students, however, her book will be of less value than Lewis Way's *Alfred Adler*, published recently in the Pelican Psychology Series.

MARGARET KNIGHT

Symbols of Transformation. Volume Five of the Collected Works of C. G. Jung. Ed. by HERBERT READ, MICHAEL FORDHAM and GERHARD ADLER; transl. by R. C. F. HULL. London: Routledge and Kegan Paul. 1956. Pp. xxix + 567. 35s.

The Analysis of Dreams. By MEDARD BOSS, transl. by ARNOLD J. POMERANS. London: Rider, 1957. Pp. 223. 25s.

Mysticism, Sacred and Profane. By R. C. ZAEHNER. Oxford: The Clarendon Press, 1957. Pp. xviii + 256. 42s.

Followers of Jung will be glad to see the fifth volume of this admirably produced series. The book is lavishly illustrated, with 64 plates and 43 text-figures, and it is sufficiently well known to need no further comment.

Prof. Boss's *Analysis of Dreams* contains an interesting Foreword by Dr E. B. Strauss, and is an attempt to approach dreams in a phenomenological and existentialist way. The author has studied Freudian and Jungian interpretations of dreams, and become discontented with both. After criticizing these and other theories of dream-interpretation he puts forward his own approach.

Prof. Zaehner explains that the purpose of his book on mysticism is to solve, perhaps presumptuously, the problem of the proper relationship between the monist's felt conviction that soul and God are one thing on the one hand, and the direct confrontation of 'I' and 'Thou' on the other. He attempts to distinguish between what seem to be radically different types of mysticism and to relate them to one another.

R. W. PICKFORD

The Child and the Family. By D. W. WINNICOTT. London: Tavistock. 1957. Pp. ix + 147. 12s. 6d.

The Child and the Outside World. By D. W. WINNICOTT. London: Tavistock. 1957. Pp. viii + 190. 16s.

These books contain a collection of papers written or broadcast during the past 18 years by Dr Winnicott. The papers' arrangement provides some continuity of ideas but does not prevent some overlapping and repetition which becomes a little tedious if the books are read at a sitting. The first volume is addressed to the parents of young children but is likely to be read and appreciated more by those professionally concerned with the welfare of children since, as the author suggests, the ordinary mother is usually too much involved in practical tasks to reflect much on theoretical aspects of family relationships. The first section is designed to give confidence to the 'ordinary devoted mother' who wishes to be allowed to get to know her baby and to begin feeding and caring for it in her own way, without constant interference from doctors and nurses. It provides a simple account of psychoanalytical views on the mother-child relationship, infant feeding, and the child's early emotional and intellectual development with the minimum of reference to the origins of these ideas. The more intelligent lay parent might well inquire what evidence exists for some of the statements. Some readers may consider the author's views of the 'ordinary devoted mother' and the natural wisdom she brings to her task as being over-optimistic but, in other respects, many of his practical implications—on, for example, feeding, weaning, and toilet training—are now fairly generally accepted by paediatricians.

The second volume is addressed to teachers and social workers concerned with the care of children not their own. One section deals with nursery school education, sex education, adoption and related problems; another with 'Some reflexions on impulse in children'. The third section concerns the problems of evacuated children during the recent war and contains important implications for the problems of separation and children's institutions. Dr Winnicott was closely concerned with the setting up of hostels for difficult evacuated children and his account of the problems encountered and the methods for dealing

with them provides much information of theoretical and practical value today. One talk on 'The deprived mother' draws attention to a side of the separation problem which is frequently neglected. While this third section is perhaps the most interesting for the psychological reader, the two books together provide a worthy response to the requests of a wider audience for the publication of Dr Winnicott's works. His style is lucid, his manner friendly, and his years of experience provide much wise insight into child behaviour and parental attitudes even for those who do not accept his theoretical bias.

AGNES E. NAUGHTON

Psychopharmacology. Ed. by NATHAN S. KLINE. Washington: American Association for the Advancement of Science. 1956. Pp. 165. 32s.

Techniques for the Study of Behavioral Effects of Drugs. By P. B. DEWS, B. F. SKINNER, and ten others. 1956. *Annals of the New York Academy of Sciences*, **65**, 247-356. 21s.

The Pharmacology of Psychotomimetic and Psychotherapeutic Drugs. By S. S. KETY, and fifty-five others. 1957. *Annals of the New York Academy of Sciences*, **66**, 417-840. 35s.

In the last few years 'psychopharmacology' has become a subject for busy research, especially in the United States. New 'tranquillizing' and 'hallucinogenic' drugs have attracted the attention of many different kinds of investigators, neurophysiologists, biochemists and psychiatrists, as well as psychologists and pharmacologists. Co-operation between them is growing because the relations between the central nervous system and the drugs which act upon it are extremely complex.

Psychopharmacology deals, in spite of its general title, almost entirely with the clinical uses of chlorpromazine and reserpine. These drugs are now widely given to make psychiatric patients more manageable. They are believed to do this by calming them without making them drowsy or inefficient, but case histories and clinical trials show that individual differences are large. The last three papers examine modes of action of several drugs, including mescaline. Lysergic acid, the 'psychotic' effects of which were discovered by accident in 1943, produces profound psychological disturbances in doses as small as 0.0001 milligram.

Techniques for the Study of Behavioral Effects of Drugs is mainly about animals, especially operant conditioning. The effects of drugs and of reinforcement schedules may mimic each other; for example, effects of a stimulant drug on the pecking rate of pigeons may resemble those of periods of *ad libitum* feeding. Operant behaviour can discriminate reliably between drugs, and this could be useful for screening new substances. In the only paper on human subjects, interesting methods are described for analysing moods. Different drugs produced characteristic moods, and group experiments showed that moods due to some drugs were more contagious than those due to others.

The Pharmacology of Psychotomimetic and Psychotherapeutic Drugs consists of 35 papers, mostly reviews and many of them highly specialized. 'A review of comparative behavioral pharmacology' refers to over 150 titles. The reviewer of 'Pharmacotherapy in psychiatry' finds that opinions of the value of drugs in psychotherapy are divided and finally declares himself 'conservatively enthusiastic' about chlorpromazine.

Several of the papers deal with lysergic acid diethylamide (LSD) and serotonin. Serotonin causes isolated smooth muscle to contract, and this effect is reduced by LSD. Since serotonin occurs naturally in the brain, this suggested that LSD might produce its artificial psychoses by interfering with it. Serotonin might thus have a role in maintaining normal mental functions, and be involved in real-life psychoses. If a normal mouse is placed facing downhill on a tiled surface it pushes backwards with its forelegs, as if to prevent itself sliding. LSD makes mice walk backwards, and this has been attributed to an 'hallucination' of sliding downhill. Injections of serotonin sometimes protect them from this symptom. Antagonism between the two substances can also be shown with other responses, such as conditioned avoidance behaviour. In other experiments, however, both on behaviour and with isolated tissues, the effects of serotonin and of LSD have been not opposite but similar. Moreover, several drugs antagonize serotonin in isolated tissues but have no psychological effects.

The common aim of all this work is to help explain mechanisms of the central nervous system and of mental illness. The findings reported are new and interesting. It is not surprising that they leave many questions unanswered, especially when so little is still known about drugs like alcohol and barbiturates which have been studied for many years. 'I feel strongly', says Prof. Kety summing up, 'that these data are trying to tell us a story, and I am happy that so many of us appear to be listening.'

HANNAH STEINBERG

OTHER PUBLICATIONS RECEIVED

- Being and Nothingness: an Essay on Phenomenological Ontology.* By JEAN-PAUL SARTRE. Transl. and introduced by H. E. BARNES. London: Methuen. 1957. Pp. lxvii + 636. 50s.
- The Tibetan Book of the Dead.* Third ed. By W. Y. EVANS-WENTZ. With a psychological commentary by C. G. JUNG. London: Oxford University Press. 1957. Pp. lxxxiv + 249. 25s.
- The Road to Inner Freedom: the Ethics.* By B. SPINOZA. Edited and introduced by D. D. RUNES. New York: Philosophical Library. 1957. Pp. 215. \$3.00.
- Personality Traits between Puberty and Adolescence: their Relationships, Development and Constancy with Reference to their Relation to School Achievement.* By R. O. VIITAMÄKI. Helsinki: Finnish Academy of Science and Letters. 1956. Pp. 183. F. mk. 700.
- The Art of Marriage.* By MARY MACAULAY. Harmondsworth: Penguin. 1957. Pp. 127. 2s. 6d.
- Problems of Human Pleasure and Behaviour.* By MICHAEL BALINT. London: Hogarth Press. 1957. Pp. 300. 25s.
- Hamlet's Mouse Trap: a Psychoanalytic Study of the Drama.* By ARTHUR WORMHOUDT. New York: Philosophical Library. 1956. Pp. 221. \$3.50.
- Truth and Meaning.* By DAVID GREENWOOD. New York: Philosophical Library. 1957. Pp. 114. \$3.75.
- Experimental Psychology and Other Essays.* By I. P. PAVLOV. New York: Philosophical Library. 1957. Pp. 653. \$7.50.
- The Nature of Conflict: Studies on the Sociological Aspects of International Tensions.* By The International Sociological Association, in collaboration with J. BERNARD, T. H. PEAR, R. ARON and R. C. ANGELL. Paris: UNESCO. London: H.M.S.O. 1957. Pp. 314. 21s.
- Neurotic Interaction in Marriage.* Ed. by VICTOR W. EISENSTEIN. London: Tavistock Publications. 1957. Pp. 352. 30s.
- The Psychopathology of Prostitution.* Rev. ed. By EDWARD GLOVER. London: Institute for the Study and Treatment of Delinquency. 1957. Pp. 20. 1s. 6d.
- The Problem of Homosexuality.* Ed. by EDWARD GLOVER. London: Institute for the Study and Treatment of Delinquency. 1957. Pp. 40. 3s.
- Family and Kinship in East London.* By MICHAEL YOUNG and PETER WILLMOTT. London: Routledge and Kegan Paul. 1957. Pp. 232. 25s.
- Patients and Doctors.* By KENNETH WALKER. Harmondsworth: Penguin. 1957. Pp. 182. 3s. 6d.
- Culture and Mental Disorders.* By RALPH LINTON. Illinois: Thomas. Oxford: Blackwell. 1956. Pp. 139. 34s.
- The Mind Goes Forth.* By HARRY and BONARO OVERSTREET. London: Peter Davies. 1957. Pp. 350. 21s.
- Psychiatric Studies.* By C. G. JUNG. Volume One of Jung's Collected Works. Transl. by R. F. C. HULL. London: Routledge and Kegan Paul. 1957. Pp. 269. 25s.
- Methods and Criteria of Reasoning: an Inquiry into the Structure of Controversy.* By RUPERT CRAWSHAY-WILLIAMS. London: Routledge and Kegan Paul. 1957. Pp. 296. 32s.

- Envy and Gratitude: a Study of Unconscious Sources.* By MELANIE KLEIN. London: Tavistock Publications. 1957. Pp. 101. 12s. 6d.
- Schizophrenia: Somatic Aspects.* Ed. by DEREK RICHTER. London: Pergamon Press. 1957. Pp. 181. 40s.
- Individual Differences in Night-Vision Efficiency.* M.R.C. Special Report No. 294. By M. H. PIRENNE, F. H. C. MARRIOTT, and E. F. O'DOHERTY. London: H.M.S.O. 1957. Pp. 83. 8s.
- The Stress of Life.* By HANS SELYE. London: Longmans, Green. 1957. Pp. 324. 18s.
- Selected Contributions to Psycho-analysis.* By JOHN RICKMAN. Ed. by W. C. M. SCOTT. London: Hogarth Press. 1957. Pp. 411. 30s.
- Disaster: a Psychological Essay.* By MARTHA WOLFENSTEIN. London: Routledge and Kegan Paul. 1957. Pp. 231. 23s.
- Psychological Services for Schools.* Ed. by W. D. WALL. Hamburg: UNESCO Institute for Education. 1956. Pp. 150. 5s.
- The Trickster: a Study in American Indian Mythology.* By PAUL RADIN. London: Routledge and Kegan Paul. 1956. Pp. 211. 21s.
- Changing Concepts of Psychoanalytic Medicine.* Ed. by S. RADO and G. E. DANIELS. New York and London: Grune and Stratton. 1956. Pp. 248. \$6.75.
- Experimental Design and its Statistical Basis.* By D. J. FINNEY. London: Cambridge University Press. 1955. Pp. 169. 30s.
- Man and Automation.* By L. LANDON GOODMAN. Harmondsworth: Penguin. 1957. Pp. 286. 3s. 6d.
- Consumer Problems and Personal Finance.* Third ed. By ARCH. W. TROELSTRUP. London and New York: McGraw-Hill. 1957. Pp. 511. 45s.
- The Psychology of Sexual Emotion: The Basis of Selective Attraction.* By VERNON W. GRANT. London: Longmans, Green. 1957. Pp. 270. 35s.
- Structural Psychology: De Humani Mentis Fabrica.* By D. and K. STANLEY-JONES. Bristol: Wright. 1957. Pp. 179. 21s.
- Nostalgia: a Psychoanalytic Study of Marcel Proust.* By MILTON L. MILLER. London: Gollancz. 1957. Pp. 306. 21s.
- The Psychology of Persuasive Speech.* Second ed. By ROBERT T. OLIVER. London: Longmans, Green. 1957. Pp. 466. 35s.
- Probability and Scientific Inference.* By G. SPENCER BROWN. London: Longmans, Green. 1957. Pp. 154. 15s.
- The Placement of Adoptive Children.* By J. RICHARD WITTENBORN. Illinois: Thomas. Oxford: Blackwell. 1957. Pp. 189. 36s.
- Mental Depressions and their Treatment.* By SAMUEL HENRY KRAINES. New York and London: The Macmillan Company. 1957. Pp. 555. 56s.
- Applicability of Factor Analysis in the Behavioral Sciences: a Methodological Study.* By STEN HENRYSSON. Stockholm Studies in Educational Psychology, No. 1. Stockholm: Almqvist and Wiksell. 1957. Pp. 156. Sw. kr. 15.00.
- Studies in Hysteria.* By JOSEF BREUER and SIGMUND FREUD. Translated and edited by J. and A. STRACHEY. London: Hogarth Press. 1957. Pp. 335. 25s.
- Essays in Linguistics.* By JOSEPH H. GREENBERG. Chicago University Press. London: Cambridge University Press. 1957. Pp. 108. 22s. 6d.

- Alcoholism: Basic Aspects and Treatment.* Ed. by HAROLD E. HIMWICH. Washington: American Association for the Advancement of Science. London: Bailey Bros. and Swinfen. 1957. Pp. 212. 52s.
- Measurement and Evaluation in Education.* By J. M. BRADFIELD and H. S. MOREDOCK. New York and London: The Macmillan Co. 1957. Pp. 509. 38s. 6d.
- Sociometric Testing: A Guide for Teachers.* By M. L. NORTHWAY and L. WELD. Toronto University Press. London: Oxford University Press. 1957. Pp. 72. 16s.
- Cultural Pluralism and the American Idea: An Essay in Social Philosophy.* By H. M. KALLEN. Pennsylvania University Press. London: Oxford University Press. 1957. Pp. 208. 40s.
- Philosophical Psychology.* By J. F. DONCEEL. London: Sheed and Ward. 1957. Pp. 363. 21s.
- The Meaning of Persons.* By P. TOURNIER. Transl. from French by E. HUDSON. London: S.C.M. Press. 1957. Pp. 238. 21s.
- The Psychologic Study of Man.* By J. MONEY. Illinois: Thomas. Oxford: Blackwell. 1957. Pp. 216. 36s.
- The Mentally Ill Child: A Guide for Parents.* By S. B. GETZ and E. L. REES. Illinois: Thomas. Oxford: Blackwell. 1957. Pp. 88. 25s.
- The Three Faces of Eve.* By C. H. THIGPEN and H. M. CLECKLEY. London: Secker and Warburg. 1957. Pp. 313. 18s.
- The Challenge of Soviet Education: The Study of Education as a Weapon.* By G. S. COUNTS. London: McGraw-Hill. 1957. Pp. 330. 45s.
- Man and Culture: An Evaluation of the Work of Malinowski.* Ed. by RAYMOND FIRTH. London: Routledge and Kegan Paul. 1957. Pp. 292. 32s.
- Statistical Theory: The Relationship of Probability, Credibility, and Error.* By LANCELOT HOGBEN. London: George Allen and Unwin. 1957. Pp. 510. 45s.
- Soviet Education.* Ed. by G. L. KLINE. London: Routledge and Kegan Paul. 1957. Pp. 192. 21s.
- Executive Performance and Leadership.* By C. L. SHARTLE. London: Staples. 1957. Pp. 302. 30s.
- Living Magic: The Realities Underlying the Psychical Practices and Beliefs of Australian Aborigines.* By RONALD ROSE. London: Chatto and Windus. 1957. Pp. 222. 18s.
- The Dynamics of Interviewing: Theory, Technique, and Cases.* By R. L. KAHN and C. F. CANNELL. New York: Wiley. London: Chapman and Hall. 1957. Pp. 368. 62s.
- Pain and Pleasure: A Study of Bodily Feelings.* By THOMAS S. SZASZ. London: Tavistock Publications. 1957. Pp. 301. 30s.
- Light, Colour and Vision.* By YVES LE GRAND. Transl. from French by R. W. G. HUNT, J. W. T. WALSH and F. R. W. HUNT. London: Chapman and Hall. 1957. Pp. 512. 63s.
- Mental Health Aspects of Urbanization.* By several authors. London: World Federation for Mental Health. London: H. K. Lewis. 1957. Pp. 46. 6s.
- Mental Health and the World Community.* Ed. by FRASER BROCKINGTON. London: World Federation for Mental Health. London: H. K. Lewis. 1957. Pp. 118. 10s.
- Sigmund Freud: Life and Work.* Vol. III. *The Last Phase, 1919-1939.* By ERNEST JONES. London: Hogarth Press. 1957. Pp. 536. 35s.

- How Children Learn to Speak.* By M. M. LEWIS. London: Harrap. 1957. Pp. 143. 8s. 6d.
- Identity and Interpersonal Competence.* By N. N. FOOTE and L. S. COTTRELL, Jr. Chicago: University Press. London: Cambridge University Press. 1956. Pp. 305. 37s. 6d.
- Closed Ranks: an Experiment in Mental Health Education.* By E. CUMMING and J. CUMMING. Harvard University Press (Commonwealth Fund). London: Oxford University Press. 1957. Pp. 192. 28s.
- Problemkinder.* By VARIOUS AUTHORS. Bremen: Institut für Jugendkunde. 1957. Pp. 64. DM. 4.20.
- The Feeling of Superiority and Anxiety-Superior: the Ottawa Trial Survey on Personality.* By E. L. REMITS. Ottawa: Runge Press. 1957. Pp. 82. No price given.
- Peter and Caroline: A Child Asks about Childbirth and Sex.* By S. HEGELER. Transl. from Danish by M. MICHAEL. London: Tavistock. 1957. Pp. 30. 7s. 6d.
- Helping Yourself with Psychiatry.* By FRANK S. CAPRIO. New York: Prentice-Hall. London: Bailey Bros. and Swinfen. 1957. Pp. 256. 40s.
- The Magic of Believing for Young People.* By CLAUDE BRISTOL. New York: Prentice-Hall. London: Bailey Bros. and Swinfen. 1957. Pp. 204. 24s.
- On the Unified Factor Theory of Mind.* By Y. AHMAVAARA. Helsinki: Finnish Academy of Science and Letters. 1957. Pp. 176. F.mk. 750.
- Psychotherapy: A Modern Theory and Practice.* By E. L. PHILLIPS. London and New York: Staples. 1957. Pp. 334. 30s.
- Homicide in American Fiction, 1798-1860: A Study in Social Values.* By D. B. DAVIS. Cornell University Press. London: Oxford University Press. 1957. Pp. 346. 30s.
- Systematic Sociology: An Introduction to the Study of Sociology.* By KARL MANNHEIM. Ed. by J. S. ERÖS and W. A. C. STEWART. London: Routledge and Kegan Paul. 1957. Pp. xxx+169. 24s.
- The Educational Attainment of Bilingual Children in Relation to their Intelligence and Linguistic Background.* By W. R. JONES, J. R. MORRISON, J. ROGERS and H. SAER. Cardiff: University of Wales Press. 1957. Pp. 51. 3s. 6d.
- The Doctor Eyes the Poor Reader.* By D. G. SCHUBERT. Illinois: Thomas. Oxford: Blackwell. 1957. Pp. 101. 27s. 6d.
- Poor Monkey: The Child in Literature.* By PETER COVENEY. London: Rockliff. 1957. Pp. 297. 30s.
- New Developments in Analytical Psychology.* By MICHAEL FORDHAM. London: Routledge and Kegan Paul. 1957. Pp. 214. 25s.
- An Introduction to Social Anthropology.* Vol. 2. By RALPH PIDDINGTON. Edinburgh: Oliver and Boyd. 1957. Pp. 819. 30s.
- Psychology and Psychological Medicine for Nurses.* By PORTIA HOLMAN. London: Heinemann. 1957. Pp. 156. 10s. 6d.
- Glory Reflected: Sigmund Freud—Man and Father.* By MARTIN FREUD. London: Angus and Robertson. 1957. Pp. 218. 21s.
- Hypnosis in Health and Sickness.* By G. AMBROSE and G. NEWBOLD. London: Staples. Pp. 196. 15s.
- Gruppen Psychotherapie.* Ed. by H. HILTMANN, K. H. WEWETZER and H. R. TEIRICE. Bern and Stuttgart: Hans Huber. 1957. Pp. 331. DM. 16.